inside

SPRING 2014
inSIDE • Vol. 12 No. 1

Innovatives Supercomputing in Deutschland

inside is published two times a year by The GAUSS Centre for Supercomputing (HLRS, LRZ, JSC)

Publishers
Prof. Dr. A. Bode | Prof. Dr. Dr. Th. Lippert | Prof. Dr.-Ing. Dr. h.c. Dr. h.c. M. M. Resch

Editor
F. Rainer Klank, HLRS klank@hlrs.de

Design
Linda Weinmann, HLRS weinmann@hlrs.de
Sebastian Zeeden, HLRS zeeden@hlrs.de
Pia Heusel, HLRS heusel@hlrs.de

Editorial

Welcome to this new issue of inside, the journal on Innovative Supercomputing in Germany published by the Gauss Centre for Supercomputing (GCS). In this issue, there is a clear focus on applications. While the race for ever faster systems is accelerated by the challenge to build an Exascale system within a reasonable power budget, the increased sustained performance allows for the solution of ever more complex problems. The range of applications presented is impressive covering a variety of different fields.

GCS has been a strong player in PRACE over the last years. In 2015 will continue to provide access to leading edge systems for all European researchers through the evaluation process of PRACE. By the end of 2015 GCS will have delivered CPU time worth 100 Million Euro. It is noteworthy that computational costs as offered by GCS are even lower than comparable Cloud offerings. The concept of dedicated centers for research is working well both for the user community and the funding agencies. For PRACE we will give a review on the results of the last call or simulation proposals.

HPC continues to be an important issue in the German research community. The working group of the Wissenschaftsrat (Science Council), which is currently evaluating the existing German funding scheme for HPC and discussing possible improvements, is expected to deliver its report in July. At the same time a number of software initiatives like the HPC Software Initiative of the Federal Ministry of Science (BMBF) or the SPPEXA initiative of the German Research Society (DFG) start to create interesting results as is shown in our project section.

With respect to hardware GCS is continuously following its roadmap. New systems will become available at HLRS in 2014 and at LRZ in 2015. System shipment for HLRS is planned to start on August 8 with a chance to start general operation as early as October 2014. We will report on further progress.

As usual, this issue includes information about events in supercomputing in Germany over the last months and gives an outlook of workshops in the field. Readers are invited to participate in these workshops which are now part of the PRACE Advanced Training Center and hence open to all European researchers.

- Prof. Dr. A. Bode (LRZ)
- Prof. Dr. Dr. Th. Lippert (JSC)
- Prof. Dr.-Ing. Dr. h.c. Dr. h.c. M. M. Resch (HLRS)

PRACE: Results of the 8th Regular Call

• Walter Nadler

Jülich Supercomputing Centre (JSC)

The Partnership for Advanced Computing in Europe (PRACE) is continuously offering supercomputing resources on the highest level (tier-0) to European researchers.

The Gauss Centre for Supercomputing (GCS) is currently dedicating shares of its IBM iDataPlex system SuperMUC in Garching, of its Cray XE6 system Hermit in Stuttgart, and of its IBM Blue Gene/Q system JUQUEEN in Jülich. France, Italy, and Spain are dedicating shares on their systems CURIE, hosted by GENCI at CEA-TGCC in Bruyères-Le-Châtel, FERMI, hosted by CINECA in Casalecchio di Reno, and MareNostrum, hosted by BSC in Barcelona.

The 8th call for proposals for computing time for the allocation time period March 4, 2014 to March 3, 2015 on the above systems closed October 15, 2013. Nine research projects have been granted a total of about 170 million compute core hours on Hermit, seven have been granted a total of about 120 million compute core hours on SuperMUC, and five proposals have been granted a total of 100 million compute core hours on JUQUEEN. One research project has been granted resources both on Hermit and SuperMUC.

Four of the newly awarded research projects are from Germany and the UK, each, three are from Italy, two are from France and Sweden, each, and one is from Belgium, Finland, the Netherlands, Spain, and Switzerland, each. The research projects awarded computing time cover again many scientific areas, from Astrophysics to Medicine and Life Sciences. More details, also on the projects granted access to the machines in France, Italy, and Spain, can be found via the PRACE web page www.prace-ri.eu/PRACE-8th-Regular-Call.

The 9th call for proposals for the allocation time period September 2, 2014 to September 1, 2015 closed March 25, 2014 and evaluation is still under way, as of this writing. The 10th call for proposals is exptected to open in September 2015.

Details on calls can be found on www.prace-ri.eu/Call-Announcements.

contact: Walter Nadler, w.nadler@fz-juelich.de

Optimizing a Seismic Wave Propagation Code for PetaScale-Simulations

The accurate numerical simulation of seismic wave propagation through a realistic three-dimensional Earth model is key to a better understanding of the Earth's interior and is used for earthquake simulation as well as hydrocarbon exploration. Seismic waves generated by an active (e.g. explosions or air-guns) or passive (e.g. earthquakes) source carry valuable information about the subsurface structure. Geophysicists interpret the recorded waveforms to create a geological model (Fig. 1). Forward simulation of waves support the survey setup assembly and hypothesis testing to expand our knowledge of complicated wave propagation phenomena. One of the most exciting applications is the full simulation of the frictional sliding during an earthquake and the excited waves that may cause damage on built infrastructure.

One of the biggest challenges today is to correctly simulate the high frequencies in the wave field spectrum. These high frequencies can be in resonance to buildings or enable a clearer image of the subsurface due to their sensitivity to smaller structures. To match reality and simulated results, high accuracy and fine resolution in combination with geometrical flexibility are the demands on modern numerical schemes. Our software package SeisSol implements a highly optimized Arbitrary high-order DERivatives Discontinuous Galerkin (ADER-DG) method that works on unstructured tetrahedral meshes. The ADER-DG algorithm is high-order accurate in space and time to ensure lowest dispersion errors for simulating waves propagating over large distances.

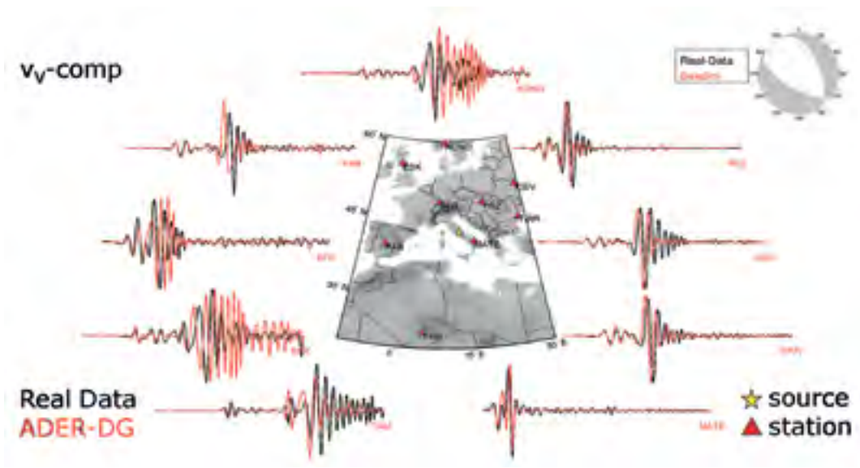


Figure 1: Waveforms of the 2009 L'Aquila earthquake (yellow star). Synthetic seismograms produced by SeisSol are depicted as red lines and compared to recorded real data (black lines) at 9 stations (red triangles). The seismograms span 800 seconds and are filtered to periods of T ≥ 33 s. Figure obtained from Wenk et al. (2013).

In a collaboration of the research groups on Geophysics at LMU Munich (Dr. Christian Pelties, Dr. Alice-Agnes Gabriel, Stefan Wenk) and HPC at TUM, Munich (Prof. Dr. Michael Bader, Alexander Breuer, Sebastian Rettenberger, Alexander Heinecke), SeisSol has been re-engineered with the explicit goal to realize petascale simulations and to prepare SeisSol for the next generation of supercomputers.

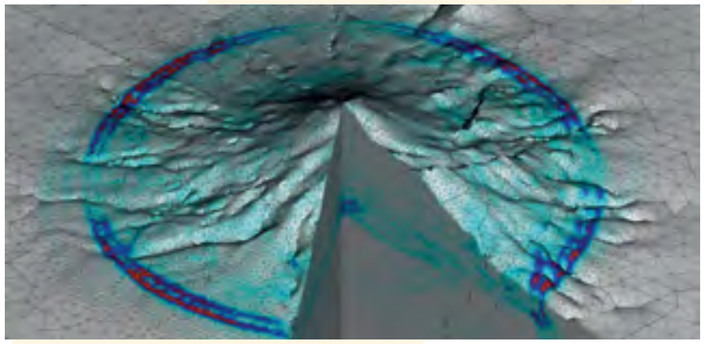


Figure 2: Tetrahedral mesh and simulated wave field (after 4 seconds simulated time) of the Mount Merapi scenario. To show the wave field in the volcano's interior, the front section is virtually removed.

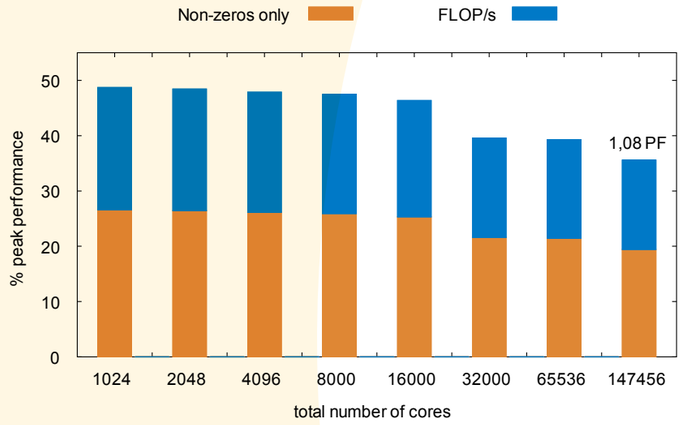


Figure 3: Strong scaling results of the 100 million cell Mount Merapi simulation. Given is the achieved percentage of peak performance (blue) and the respective fraction of non-zero floating point operations (orange). While reaching almost 50% of peak performance on 1,024 cores, a peak efficiency of 35% is maintained on all 147,456 cores of SuperMUC reaching 1.08 Petaflop/s. The observed performance decrease from 16,000 to 32,000 cores is caused by the slower inter-island communication on SuperMUC.

Optimizing SeisSol for PetaScale-Simulations

With this goal in mind, three key performance and scalability issues were addressed in SeisSol:

- Optimized Matrix Kernels:

In SeisSol's ADER-DG method, the element-local numerical operations are implemented via matrix multiplications. For these operations, highly optimized, hardware-aware kernels were generated by a custom-made code generator. Depending on the sparsity pattern of the matrices and on the achieved time to solution, sparse or dense kernels were selected for each individual operation. Due to more efficient exploitation of current CPUs' vector computing features, speedups of 5 and higher were achieved.

- Hybrid MPI+OpenMP-Parallelization:

SeisSol's parallelization was carefully changed towards a hybrid MPI+OpenMP approach. The resulting reduction of MPI processes had a positive effect on parallel efficiency especially in setups with huge numbers of compute cores. Even communication and management routines were parallelized, as these turned into performance obstacles on large numbers of cores.

- Parallel Input of Meshes:

SeisSol's key strength is the accurate modelling of highly complicated geometries via unstructured adaptive tetrahedral meshes. These are generated by an external mesh generator (SimModeler by Simmetrix). Thus, another crucial step towards petascale simulations was the use of netCDF and MPI I/O to input such large meshes. The new mesh reader can handle meshes with more than a billion grid cells, and scales on all 147k cores of SuperMUC.

1 PetaFlop/s Sustained Performance

In cooperation with the Leibniz Supercomputing Centre, several benchmark runs were executed with SeisSol on SuperMUC. In an Extreme Scaling Period end of January 2014, a weak-scaling test with 60,000 grid cells per core was conducted: SeisSol achieved a performance of 1.4 PetaFlop/s computing a problem with nearly 9 billion grid cells and over 1,012 degrees of freedom. However, as the true highlight a simulation under production conditions was performed. We modelled seismic wave propagation in the topographically complex volcano Mount Merapi (Fig. 2) discretized by a mesh with close to 100 million cells. In 3 hours of computing time the first 5 seconds of this process were simulated. During this simulation SeisSol achieved a sustained performance of 1.08 PetaFlop/s (equivalent to 35% of peak performance, see the performance plot in Fig. 3), thus entering a performance region that usually requires noticeably larger machines.

Outlook

Details on the optimization of SeisSol, and probably results from further large-scale simulations, will be presented at this year's International Supercomputing Conference (ISC'14) in Leipzig. New challenges to tackle include effective load balancing for simulations with local time stepping and novel accelerator technologies such as Intel MIC.

Acknowledgements

We especially thank Dr. Martin Käser for pioneering the development of SeisSol and for initiating the collaborative project ASCETE. Furthermore, we thank the entire LRZ team (including our co-developer Dr. Gilbert Brietzke) for their strong support in the execution of the SuperMUC runs, and Mikhail Smelyanskiy (Intel Labs) for his support in hardware-aware optimization. Financial support was provided by DFG, KONWIHR and Volkswagen Foundation (project ASCETE).

References

[1] **Breuer, A., Heinecke, A., Rettenberger, S., Bader, M., Gabriel, A.-A., Pelties, C.**
Sustained Petascale Performance of Seismic Simulations with SeisSol on SuperMUC, International Supercomputing Conference ISC'14, accepted.

[2] **Breuer, A., Heinecke, A., Bader, M., Pelties, C.**
Accelerating SeisSol by Generating Vectorized Code for Sparse Matrix Operators, Parallel Computing–Accelerating Computational Science and Engineering (CSE), Advances in Parallel Computing 25, IOS Press, 2014.

[3] **Pelties, C., de la Punte, J., Ampuero, J.-P., Brietzke, G.B., Käser, M.**
Three-dimensional dynamic rupture simulation with a high-order discontinuous Galerkin method on unstructured tetrahedral meshes, J. Geophys. Res., 117, B02309, 2012, http://dx.doi.org/10.1029/2011JB008857.

[4] **Wenk, S., Pelties, C., Igel, H., Käser, M.**
Regional wave propagation using the discontinuous Galerkin method, Solid Earth, 4, 43-57, 2013, doi:10.5194/se-4-43-2013.

Links

http://www.ascete.de
http://seissol.geophysik.uni-muenchen.de
http://www.lrz.de

contact: Michael Bader,
bader@in.tum.de

- Michael Bader[1]
- Christian Pelties[2]

[1] TU Munich, Institut of Informatics

[2] Department of Earth and Environmental Sciences, Geophysics, University of Munich

Golden Spike Award by the HLRS Steering Committee in 2013

Industrial Turbulence Simulations at Large Scale

Simulations that are done during industrial design process have to adhere to strict time constraints in the development cycle. Thus, the limited time is often the driving factor, limiting the accuracy of such simulations. Yet, there is a strong desire to obtain as detailed and accurate simulations as possible early on, to save later development iterations. Accurate but fast simulation results can be achieved by exploiting HPC systems for these "Design of Experiment" studies.

As an example for a demanding simulation, we have a look at turbulent flows in technical devices with complex geometry. Such use cases are challenging, as the complexity of the setup has to be deployed on parallel and distributed computing resources.

not suitable in the given situations, especially when capturing highly instationary processes. Therefore, a more accurate description of turbulence is needed in these cases. The most accurate description of turbulence would be achieved by a direct numerical simulation (DNS). However, a DNS results in extreme computational loads as the smallest scales have to be resolved. Even on today supercomputing systems, DNS of complex test-cases with large domains can not be captured in a timeframe that would suit an industrial development process. A middle ground is offered by Large Eddy Simulations (LES). LES resolves the large turbulent structures directly and employs a modelling for those structures, which are too small for a viable resolution. Thus, LES reduces modeling assumptions

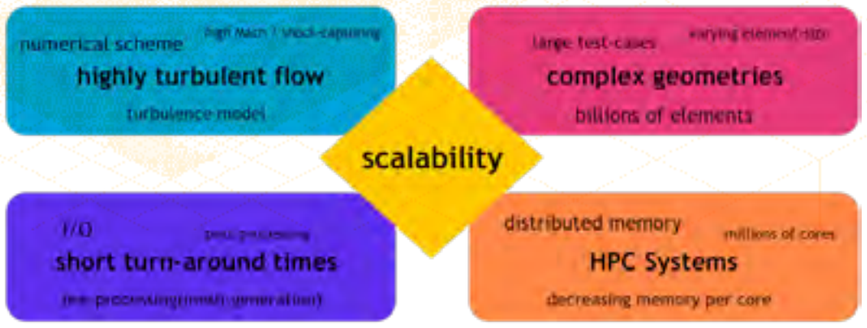


Figure 1: Challenges for industrial test-cases are highlighted. Each box is a complex topic on its own but for efficient usability everything has to be considered under the aspect of scalability.

While current commercial simulation software often employs simple turbulence models like the Reynolds-Averaged Navier-Stokes equations (RANS) to reduce the simulation time (and meshing overhead) these models are not suitable for the turbulence to a minimum, while enabling the simulation of relevant geometrical objects.

In order to achieve suitable runtimes when using the considerably more expensive LES model instead of RANS, the simulations need to be run on more processes. Yet, commercial solvers currently do not scale very well and therefore are of limited use in large simulations. Often, solvers stop scaling in a range below the average cluster size in large industries development departments. Therefore, highly scalable academic solvers are investigated and tuned to tackle industrial use-cases.

Academic codes that are used for both research and education evolve over years. With changing developers over the years and functionality-oriented feature implementation, a versatile solver is created. Optimizations of such codes aim for massively parallel execution on distributed memory systems. With the help of parallel performance tools like Vampir and code-analysis tools like the Valgrind tool suite it is possible to remove performance bottlenecks and non-scaling memory problems from the code. Especially the usage of MPI-collectives instead of less efficient isend/irecv implementations improve both, scalability and memory demand on Infiniband based clusters.

Besides code optimizations it is clear that some parts of the code need to be redesigned to achieve higher levels of distributed parallelism. Especially the interface between the mesh generation tool and the solver impose problems in general unstructured solvers. Usually the solver reads in proprietary formats from different mesh generation tools such as ICEM or Gambit, and uses a partitioning method like ParMetis for mesh partitioning. However, the process of determining the neighborhood relations between the elements and between the different processes lead to

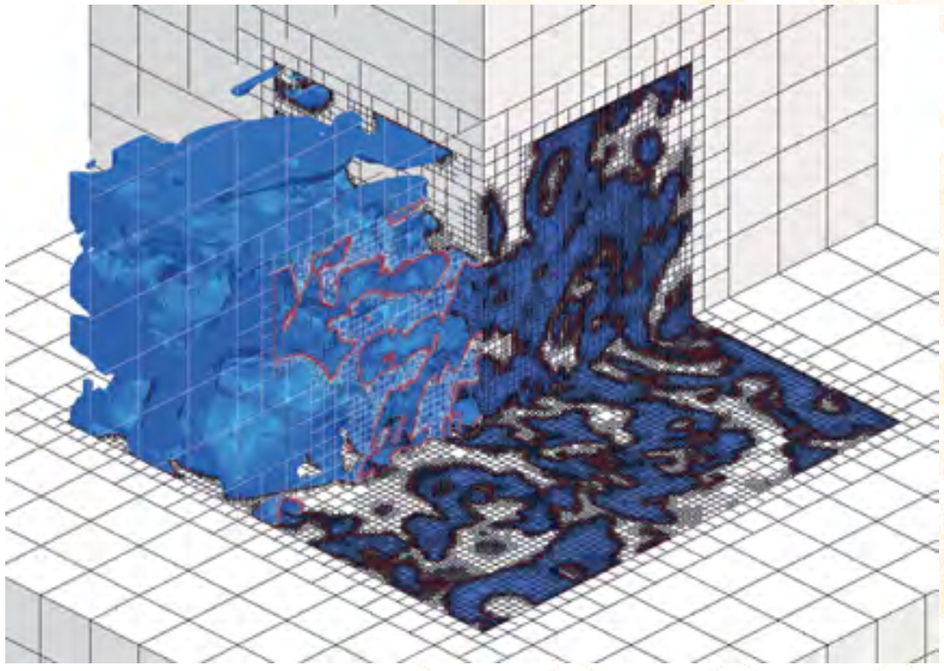


Figure 2: A discretized porous medium inside a fluid domain. (Solid elements are depicted in blue, elements with boundary conditions are represented in red, fluid elements are white.)

scalability problems early in the execution of the solver.

To efficiently handle this mesh initialization step, an approach suggested by Klimach [1] can be adopted. The proprietary mesh format of mesh generation tools is converted by a pre-processing tool, providing a file-format including neighborhood information and allowing simple distributed reading. In this conversion the mesh is read in and the individual elements are mapped into an order, which preserves physical locality of the elements using a space-filling curve. This allows a suitable partitioning for arbitrary number of processes by simply cutting the list of elements into non-overlapping chunks without any knowledge of the entire mesh topology. I.e. the initialization of the simulation run is fully parallel read right from the start.

This file format now offers a very flexible and scalable handling of meshes in the solver itself. Due to the already established neighborhood relationship, each process can locally read in only its own part of the mesh, and at the same time determine the position of direct neighbors of each local element, i.e. determine the process IDs of the communication partners for each domain without global communication. Reading the mesh with this method becomes orders of magnitude faster in the solver and unnecessary memory allocations early during the solver execution are avoided.

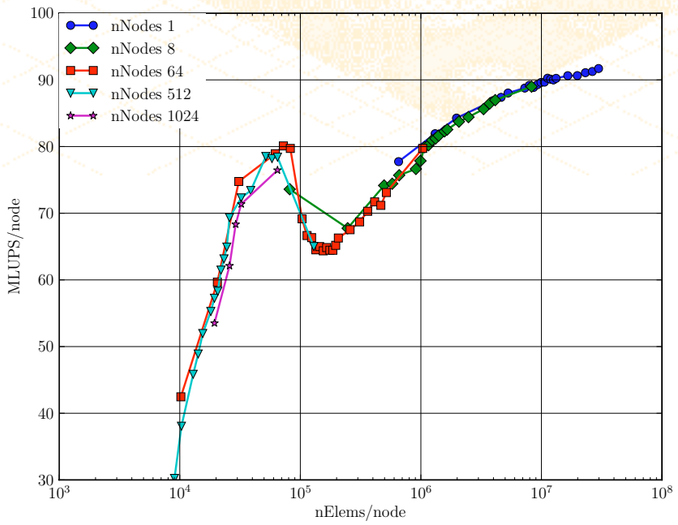


Figure 3: Performance map for the Lattice Boltzmann solver MUSUBI.

Parallel Mesh Generation

The described two-step mesh preparation is an example of how the problems of serial parts in a simulation tool-chain can affect overall running times. The inherently serial nature of e.g. commercial mesh-generation tools poses a two-fold problem when thinking in both distributed memory and parallel I/O. The conversion of the format in this case eases the problem on the solver side, however the problem still remains in a different form in the pre-processing step. On the one hand the generation of the actual mesh itself is still serial and might therefore be limited in element count due to memory and on the other hand the conversion still needs to be performed whenever the geometry is changed or the mesh refined and adapted. A more suitable solution in face of the growing parallelism is the parallel generation of meshes directly in a suitable format.

For this we develop a parallel mesh generation tool based on an octree mesh representation. This tool, called Seeder [2], provides automated parallel mesh generation in the context of the end-to-end parallel APES [3] simulation suite. It directly generates and stores the mesh-data in a format suitable for parallel processing on distributed systems. The main performance measure for mesh generation is the size of the meshes it can produce. Therefore memory scalability is more important than run-time scalability. The largest mesh that was obtained on Hermit at HLRS consists of 66 billion cubes. A mesh of this size requires a total of 4k processes. To run a simulation on this mesh using the Lattice Boltzmann solver Musubi, the solver needs to run on nearly 100k processes. With a serial mesh generation the generation of such a mesh would not be possible and therefore the usage of the entire supercomputer not efficiently possible.

Summary

Academic solvers are capable of reducing runtime considerably in product-development by the efficient usage of HPC systems. But a complete parallel tool-chain is needed, including (memory-) scalable mesh generation and solver initialization steps, in order to be able to also efficiently use the HPC systems from tomorrow for larger, more precise and/or faster simulations.

References

[1] **Klimach, H. and Roller, S.**
Distributed coupling for multiscale simulations. In P. Ivanyi and B. Topping, editors, Proceedings of the Second International Conference on Parallel, Distributed, Grid and Cloud Computing for Engineering. Civil-Comp Ltd., 2011.

[2] **Harlacher, D.F., Hasert, M., Klimach, H., Zimny, S., Roller, S.**
Tree based voxelization of stl Data. In Michael Resch, Xin Wang, Wolfgang Bez, Erich Focht, Hiroaki Kobayashi, and Sabine Roller, editors, High Performance Computing on Vector Systems 2011, pages 81{92. Springer Berlin Heidelberg, 10.1007/978-3-642-22244-3 6. 2012.

[3] **Roller, S., Bernsdorf, J., Klimach, H., Hasert, M., Harlacher, D., Cakircali, M., Zimny, S., Masilamani, K., Didinger, L., Zudrop, J.**
An adaptable simulation framework based on a linearized octree. In Michael Resch, Xin Wang, Wolfgang Bez, Erich Focht, Hiroaki Kobayashi, and Sabine Roller, editors, High Performance Computing on Vector Systems 2011, pages 93{105. Springer Berlin Heidelberg, 10.1007/978-3-642-22244-3 7. 2012.

contact: Daniel F. Harlacher,
daniel.harlacher@uni-siegen.de

- Daniel F. Harlacher
- Harald Klimach
- Sabine Roller

University of Siegen

Applications

Golden Spike Award by the HLRS Steering Committee in 2013

High-Resolution Climate Predictions and Short-Range Forecasts

to Improve the Process Understanding and the Representation of Land-Surface Interactions in the WRF Model in Southwest Germany (WRFCLIM)

The application of numerical modelling for climate projections is an important task in scientific research since these projections are the most promising means to gain insight in possible future climate changes. During the last 2 decades, various regional climate models (RCM) have been developed and applied for simulating the present and future climate of Europe at a grid resolution of 25 to 50 km. It was found that these models were able to reproduce the pattern of temperature distributions reasonably well but a large variability was found with respect to the simulation of precipitation. The performance of the RCMs was strongly dependent on the quality of the boundary forcing, namely if precipitation was due to large-scale synoptic events. Additionally, summertime precipitation was subject of significant systematic errors as models with coarse grid resolution have difficulties to simulate convective events. This resulted in deficiencies with respect to simulations of the spatial distribution and the diurnal cycle of precipitation. The 50 km resolution RCMs were hardly capable of simulating the statistics of extreme events such as flash floods.

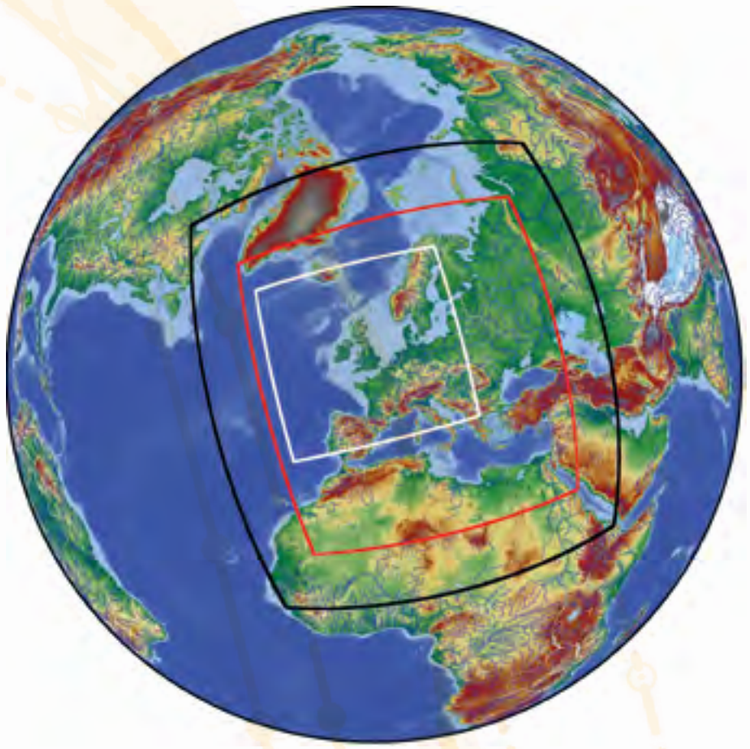


Figure 1: Domain of WRF for CORDEX Europe with 0.33° (black frame), 0.11° (red frame) and 0.0367° (white frame) resolution.

Due to these reasons, RCMs with higher grid resolution of 10-20 km were developed and extensively verified. While still inaccuracies of the coarse forcing data were transferred to the results, these simulations indicated a gain from high resolution due to better resolution of orographic effects. These include an improved simulation of the spatial distribution of precipitation and wet-day frequency and extreme values of precipitation. However, three major systematic errors remained: the windward-lee effect (i.e. too much precipitation at the windward side of mountain ranges and too little precipitation on their lee side), phase errors in the diurnal cycle of precipitation and precipitation return periods. In order to provide high-resolution ensembles and comparisons of regional climate simulations in the future, the World Climate Research Program initiated the COordinated Regional climate Downscaling EXperiment (CORDEX).

The scope of this WRFCLIM project at HLRS, is to investigate in detail the performance of regional climate simulations with WRF in the frame of EURO-CORDEX[1] at 12 km down to simulations at the convection permitting scale e.g. within the DFG funded Research Unit 1695 “Agricultural Landscapes under Climate Change–Processes and Feedbacks on a Regional Scale”[2].

Within this context the WRFCLIM objectives are to:

- provide high resolution climatological data to the scientific community

- support the quality assessment and interpretation of the regional climate projections for Europe through the contribution to the climate model ensemble for Europe (EURO-CORDEX) for the next IPCC (International Panel of Climatic Change) report.

The main objectives of the convection permitting simulations of WRFCLIM are as follows:

- Replace the convection parameterization by the dynamical simulation of the convection chain to better resolve the processes of the specific location and actual weather situation physically.

- Gain of an improved spatial distribution and diurnal cycle of precipitation through better land-surface-atmosphere feedback simulation and a more realistic representation of orography to

[1] CORDEX for the European Domain: www.euro-cordex.net

[2] An objective of this research unit is the development and verification of a convection-permitting regional climate model system based on WRF-NOAH including an advanced representation of land-surface-vegetation-atmosphere feedback processes with emphasis on the water and energy cycling between croplands, atmospheric boundary layer and the free atmosphere.

support the interpretation of the 12 km climate simulations for local applications e.g. in hydrological and agricultural management.

Special attention will be paid to the land-surface-vegetation-atmosphere feedback processes. Further, the model will be validated with high-resolution case studies applying advanced data assimilation to improve the process understanding over a wide range of temporal scales. This will also address whether the model is able to reasonably represent extreme events. In the following on the current status of regional climate simulations and land-atmosphere feedback in convection permitting simulations in WRFCLIM is reported. Further details are given e.g. in Warrach-Sagi et al. (2013a).

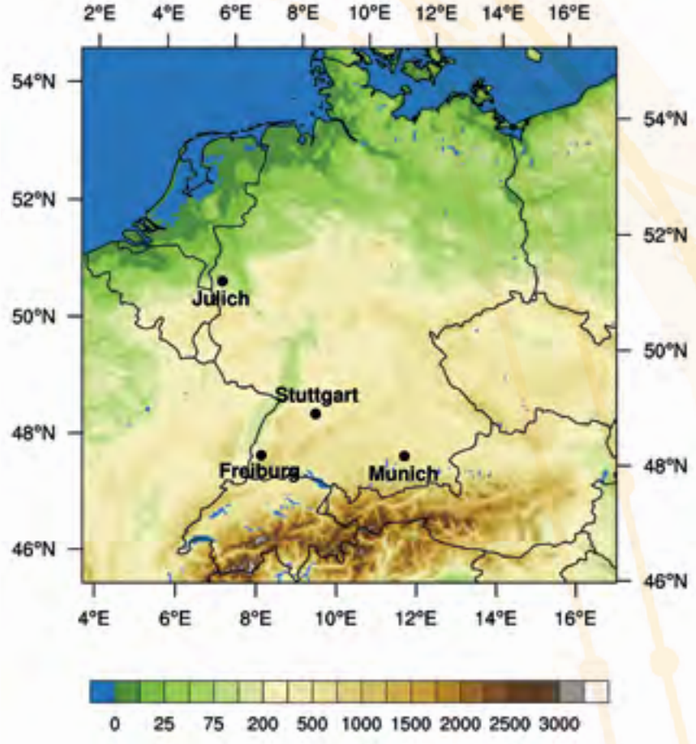


Figure 2: Domain of the second and third case study on parameterization schemes

Regional Climate Simulations

A validation run for Europe was performed with the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008) for a 22-year period (1989-2009) on the NEC Nehalem Cluster at HLRS at a grid resolution of approx. 12 km with CORDEX simulation requirements. The WRF model had been applied to Europe on a rotated latitude-longitude grid with a horizontal resolution of 0.11° and with 50 vertical layers up to 20 hPa. The model domain (red frame in Fig. 1) covers the area specified in CORDEX. The most recent reanalyses data ERA-interim of the European Center for Medium range Weather Forecast (ECMWF) is available at approx. 0.75°.

These data were used to force WRF at the lateral boundaries of the model domain. WRF was applied, one-way nested, in a double nesting approach on 0.33° (~37 km) (black frame in Fig. 1) and 0.11° (~12 km). In an additional experiment for summer 2007, when the Convective and Orographically-induced Precipitation Study (COPS) (Wulfmeyer et al. 2011) took place, a third domain with 0.0367° (~ 4 km) resolution was nested into the 0.11° domain (white frame in Fig. 1), and a convection permitting simulation was performed.

WRF is coded in Fortran and was compiled with PGI 9.04 with openMPI libraries. At 0.11° the EURO-CORDEX model domain covered 450*450*54 (in total 10,935,000) grid boxes (Fig. 1, red box) and the model was run with a time step of 60s from 1989 to 2009. For the simulation 20 nodes were requested (160 cores), the raw output data was about 80TB. Within 24 hours on the NEC Nehalem Cluster it was possible to simulate 2.5 months. The whole simulation (including waiting time between restarts) took ~8 months and ~800,000 CPU hours. For the convection permitting simulation experiment a domain of 800*800*54 (in total 34,560,000) grid boxes of approx. 4 km horizontal resolution (white frame in Fig. 1) was nested in the EURO-CORDEX model domain. A simulation with a model time step of 20s was run from Mai 15, to September 1, 2007. The simulation took one month on the NEC Nehalem Cluster. Precipitation and soil moisture results of this simulation were analyzed (Warrach-Sagi et al. 2013b, Greve et al., 2013) and due to the results in the beginning of 2012 the simulation was repeated with an updated version of WRF on the Cray XE6. WRF was compiled with PGI 11 with openMPI 1.4. This simulation was carried out for the 0.11° domain (red frame in Fig. 1) without nesting to an intermediate grid but a 30 grid box wide boundary. The simulation of the 480*468*54 (in total 12,130,560) grid boxes with a 60 s model time step was run from 1987 to 2009 on 40 nodes (1,280 cores), within 24 hours 5 months were simulated. Within 3 months the simulation was completed and approx. 90 TB of raw output data were produced. This simulation is currently under evaluation e.g. within the EURO-CORDEX ensemble of regional climate models (e.g. Vautard et al., 2013; Kotlarski et al., 2014).

Land-Atmosphere Feedback in Convection Permitting Simulations

When setting up a model for a particular region, the foremost issue is the determination of the most appropriate model configuration. Different regions experience varied conditions, and the optional setup of a model is spatially and temporarily dependent. One of the most important aspects in configuring a model is to select the parameterizations to be used and therefore sensitivity tests are an unavoidable part in improving the model performance. WRF offers a choice of various physical parameterizations to describe subgrid processes like e.g. cloud formation and radiation transport. Warrach-Sagi et al. (2013b) showed that convection permitting climate simulations with WRF bear the potential for better location of precipitation events namely in orographic terrain as it is needed by impact modelers like hydrologists and agricultural scientists. To set up the most advanced configuration for convection permitting climate simulations with WRF by the University of Hohenheim to be carried out on the Cray XE6 at HLRS in the near future the following study is performed in WRFCLIM: Experiments with various boundary layer parameterizations, radiation schemes and two land surface models are performed on 2 km resolution (convection permitting scale) to further downscale the climate data for applications in hydrology and agriculture and an improved representation of feedback processes between the land surface and the atmosphere. The first sensitivity study was performed with 54 vertical levels and a horizontal grid spacing of 2 km for a domain of 501*501 (in total 13,554,054) grid boxes with 1,088 cores on the Cray XE6 for a week in September 2009, when atmospheric observations were available from an experiment.

In a case study on the domain of Germany (Fig. 2) various combinations of 4 boundary layer schemes and 2 land surface models, including different combinations of switches in one

of them, were analyzed. The resulting ensemble of 44 simulations was compared against measurements of absolute humidity from the DIAL (Differential Absorption Lidar (light detection and ranging)) of the University of Hohenheim (Behrendt et al., 2009) close to Jülich in Western Germany in September 2009 which were taken during a campaign of the DFG Transregio 32 "Patterns in Soil-Vegetation-Atmosphere-Systems" project.

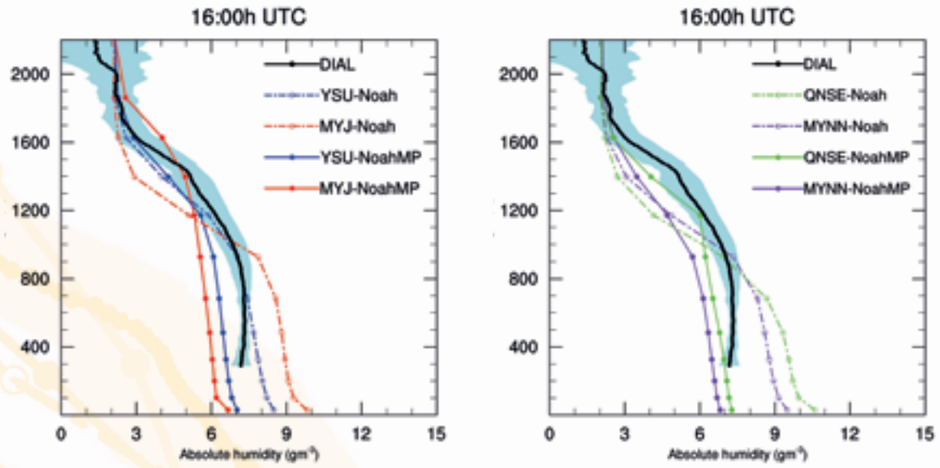


Figure 3: Vertical profiles of absolute humidity on September 8, 2009 at 16:00 UTC. Black lines represent the mean of the scanning DIAL measurements within a 2 km sector, blue shades indicate the 1-d variance of the DIAL data within the scan. Colored lines indicate different boundary layer schemes. Solid lines denote the NOAH land-surface model, while the NOAH-MP land surface model is represented in dashed lines.

It is important to be able to accurately simulate the vertical humidity and temperature profiles in the boundary layer for convection, cloud development and precipitation. Both, the boundary layer schemes and land surface models, impact the simulated vertical absolute humidity profiles. The impact of the land surface models extends also to higher altitudes, even up to entrainment zone at the top of the boundary layer (Fig. 3). Results also showed that the WRF model is sensitive on the combination of land surface model and boundary layer scheme. The comparison of model data with the high resolution lidar measurements is a first step in this ongoing research. Currently more case studies and comparisons with scanning lidar measurements are performed for different weather situations in different seasons. For this, data of field experiments in the frame of different German research projects are collected, i.e., the DFG Transregio 32 and Research Unit 1695, and the BMBF project "High definition clouds and precipitation for advancing climate prediction" HD(CP)2.

Summary of Preliminary Results

At HLRS within WRFCLIM two climate simulations with WRF were carried out for 1989-2009 to evaluate the model performance within the frame of EURO-CORDEX. The first simulation (WRF 3.1.0 on the NEC Nehalem Cluster) was evaluated by Warrach-Sagi et al. (2013b) with respect to precipitation and Greve et al. (2013) with respect to soil moisture. During all seasons, this dynamical downscaling reproduced the observed spatial structure distribution of the precipitation fields. However, a wet bias was remaining. At 12 km model resolution, WRF showed typical systematic errors of RCMs in orographic terrain such as the windward-lee effect. In the convection permitting resolution (4 km) case study in summer 2007, this error vanished and the spatial precipitation patterns further improved. This result indicates the high value of regional climate simulations on the convection-permitting scale. Soil moisture depends on and affects the energy flux partitioning at the land-atmosphere interface. The latent heat flux is limited by the root zone soil moisture and solar radiation. When compared with the in situ soil moisture observations in Southern France, where a soil moisture network was available during 2 years of the study period, WRF generally reproduces the annual cycle. The spatial patterns and temporal variability of the seasonal mean soil moisture from the WRF simulation corresponds well with two reanalyses products, while their absolute values differ significantly, especially at the regional scale. Based on these results the updated version WRF 3.3.1 was applied at Cray XE6 for EURO-CORDEX. The benefits from downscaling the reanalyses ERA-Interim data with WRF become especially evident in the probability distribution of daily precipitation data. Kotlarski et al. (2014) show that WRF 3.3.1 generally simulates the seasonal precipitation and temperature in central Europe well, namely in Germany, and fits well in the EURO-CORDEX ensemble. The evaluation within the EURO-CORDEX ensemble is still ongoing. The data is now available and used e.g. within the DFG funded Research Unit "Agricultural Landscapes under Climate Change–Processes and Feedbacks on a Regional Scale" for impact studies. However, the weakness of the 12 km resolution concerning the windward-lee effect in precipitation remains. So it is essential to proceed towards convection permitting simulations as in WRFCLIM.
The first sensitivity study of the WRF to the boundary-layer and land-surface parameterizations and comparisons with Differential Absorption Lidar for September 2009 showed significant sensitivity of WRF to the choice of the land surface model and planetary boundary layer parameterizations. The model sensitivity to the land surface model is evident not only in the lower PBL, but it extends up to the PBL entrainment zone and even up to the lower free troposphere. This study is ongoing for more experiments in spring 2013 under other weather conditions to understand the processes, their modelling and to set up a sophisticated simulation with WRF for Germany on the convection permitting scale also for climate simulations.

Acknowledgements

This work is part of the Project PAK 346/RU 1695 funded by DFG and supported by a grant from the Ministry of Science, Research and Arts of Baden-Württemberg (AZ Zu 33-721.3-2) and the Helmholtz Centre for Environmental Research - UFZ, Leipzig (WESS project). DIAL measurements were performed within the Project Transregio 32, also funded by DFG. Model simulations were

Applications

carried out at HLRS on the Cray XE6 and the NEC Nehalem Cluster within WRFCLIM and we thank the staff for their support.

References

[1] **Behrendt, A., Wulfmeyer, V., Riede, A., Wagner, G., Pal, S., Bauer, H., Radlach, M., Späth, F.**
3-Dimensional observations of atmospheric humidity with a scanning differential absorption lidar. In Richard H. Picard, Klaus Schäfer, Adolfo Comeron et al. (Eds.), Remote Sensing of Clouds and the Atmosphere XIV, SPIE Conference Proceeding Vol. 7475, ISBN: 9780819477804, 2009, Art. No. 74750L, DOI:10.1117/12.835143, 2009.

[2] **Greve, P., Warrach-Sagi, K., Wulfmeyer, V.**
Evaluating Soil Water Content in a WRF-Noah Downscaling Experiment, J. Appl. Meteor. Climatol. 52: 2312–2327, 2013.

[3] **Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., Wulfmeyer, V.**
Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. Submitted to Geoscientific Model Development, 2014.

[4] **Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W., Powers, J.G.**
A description of the Advanced Research WRF version 3. NCAR Tech Note, TN-475+STR, 113pp, 2008.

[5] **Vautard, R., Gobiet, A., Jacob, D., Belda, M., Colette, A., Deque, M., Fernandez, J., Garcia-Diez, M., Goergen, K., Guettler, I., Halenka, T., Keuler, K., Kotlarski, S., Nikulin, G., Patarcic, M., Suklitsch, M., Teichmann, C., Warrach-Sagi, K., Wulfmeyer V., Yiou, P.**
The simulation of European heat waves from an ensemble of regional climate models w ithin the EURO-CORDEX project, Climate Dynamics, 10.1007/s00382-013-1714-z, 2013.

[6] **Warrach-Sagi, K., Bauer, H.-S., Branch, O., Milovac, J., Schwitalla, T., Wulfmeyer, V.**
High-resolution climate predictions and short-range forecasts to improve the process understanding and the representation of land-surface interactions in the WRF model in Southwest Germany (WRFCLIM). In: "High Performance Computing in Science and Engineering 13", Eds: Wolfgang E. Nagel, Dietmar H. Kroener, Michael M. Resch, Springer, 529-542, 2013a.

[7] **Warrach-Sagi, K., Schwitalla, T., Wulfmeyer, V., Bauer, H.-S.**
Evaluation of a CORDEX-Europe simulation with WRF: precipitation in Germany, Climate Dynamics, DOI 10.1007/s00382-013-1727-7, 2013b.

[8] **Wulfmeyer, V., Behrendt, A., Kottmeier, C., Corsmeier, U., Barthlott, C., Craig, G.C., Hagen, M., Althausen, D., Aoshima, F., Arpagaus, M., Bauer, H.-S., Bennett, L., Blyth, A., Brandau, C., Champollion, C., Crewell, S., Dick, G., Di Girolamo, P., Dorninger, M., Dufournet, Y., Eigenmann, R., Engelmann, R., Flamant, C., Foken, T., Gorgas, T., Grzeschik, M., Handwerker, J., Hauck, C., Höller, H., Junkermann, W., Kalthoff, N., Kiemle, C., Klink, S., König, M., Krauss, L., Long, C.N., Madonna, F., Mobbs, S., Neininger, B., Pal, S., Peters, G., Pigeon, G., Richard, E., Rotach, M.W., Russchenberg, H., Schwitalla, T., Smith, V., Steinacker, R., Trentmann, J., Turner, D.D., van Baelen, J., Vogt, S., Volkert, H., Weckwerth, T., Wernli, H., Wieser, A., Wirth, M.**
The Convective and Orographically Induced Precipitation Study (COPS): The Scientific Strategy, the Field Phase, and First Highlights. COPS Special Issue of the Q. J. R. Meteorol. Soc. 137, 3-30, DOI:10.1002/qj.752, 2011.

- Kirsten Warrach-Sagi
- Josipa Milovac
- Hans-Stefan Bauer
- Andreas Behrendt
- Thomas Schwitalla
- Florian Späth
- Volker Wulfmeyer

Institute of Physics and Meteorology, University of Hohenheim

contact: Kirsten Warrach-Sagi, kirsten.warrach-sagi@uni-hohenheim.de

Golden Spike Award by the HLRS Steering Committee in 2013

A Highly Scalable Parallel LU Decomposition for the Finite Element Method

Numerical simulations with the Finite Element Method (FEM) play an important role in the current science to solve partial differential equations on a domain Ω. The equations are formulated into their equivalent weak formulation, which leads to an infinite dimensional problem. By dividing the domain Ω into subdomains and discretizing the weak formulation by a subspace with piecewise polynomial functions, we finally obtain a linear equation Au = b, where A is a sparse matrix and u the solution.

Figure 1: domain Ω on 8 processors

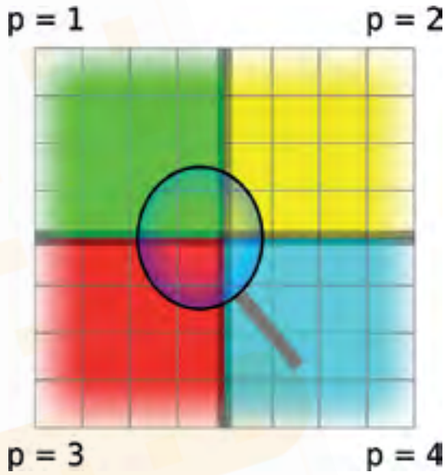


Figure 2: Cells near some interfaces

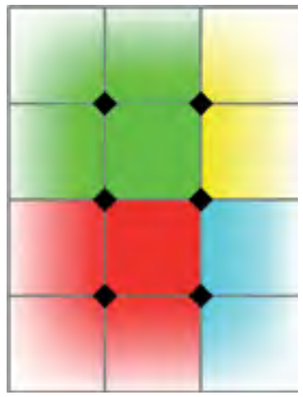


Figure 3: Nodal points

To attain a reasonable approximation of the continuous solution it is essential to use a fine meshsize for the subdomains. This can result into several million or billion unknowns. To obtain a reasonable computing time to solve the problem is indefinite, unsymmetric, or ill-conditioned. Thus, we solve the equation Au = b with the help of a parallel LU decomposition [1]. For further information about the parallel programming model, see [2].

The Idea Behind the Decomposition

Our concept of the parallel LU decomposition for matrices resulting from finite element problems is based on a nested dissection approach, see [1,3].

Figure 4: Interfaces

On P = 2^S processors, the algorithm uses S+1 steps in which sets of processors are combined together consistently and problems between these sets are solved, beginning with one processor in the first step.

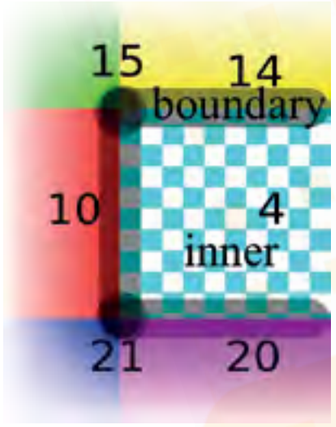


Figure 5: Inner and boundary elements in step 1.

Let us consider a domain Ω in 2D, distributed on 8 processors. The domain is discretized into cells c, and every cell is given to one processor. The collection of all cells on one processor defines a local domain Ω^p.

Every cell has some indices (or nodal points) for the finite element method. Those indices are given either on only one processor or on an interconnection between two or more processors, e.g. the crosspoint in Fig. 2 is given on processors 1, 2, 3 and 4. By defining a processor set π for each nodal point (here, e.g. $\pi(i) = \{1,2,3,4\}$) we can combine all indices with a common processor set to one interface. We have in total 21 different interfaces, cf. Table 1. Those elements with only one entry in their processor set define the inner domain of each local domain Ω^p. All other elements define a boundary part of the local domains (cf. Fig. 4).

The main observation is that all indices in the inner domain on processor p are independent (within the meaning of the finite element method) of the indices in the inner domain of any other processor q ≠ p. We can "eliminate" the inner elements (here: k=1,...,8) on every processor in parallel within an LU decomposition. For the elements on the boundary, we build the current Schur complement, which is distributed on the processors. In Fig. 5 Illustration we have an example for the inner (π_4) and boundary elements (π_{10}, π_{14}, π_{15}, π_{20}, π_{21}) on processor 4.

step 1			step 2			step 3			step 4		
k	π_k	t	k	π_k	t	k	π_k	t	k	π_k	t
1	1	1	9	1,2	1	13	1,3	1	19	3,5	1
2	2	2	10	3,4	2	14	2,4	1	20	4,6	1
3	3	3	11	5,6	3	15	1,2,3,4	1	21	3,4,5,6	1
4	4	4	12	7,8	4	16	5,7	2			
5	5	5				17	6,8	2			
6	6	6				18	5,6,7,8	2			
7	7	7									
8	8	8									

Table 1: Sorted list of interfaces with cluster t (cf. Figure 4).

For the next step, we combine each processor with another one to a new processor set, such that we now have four clusters each with two processors. The combined processor set defines

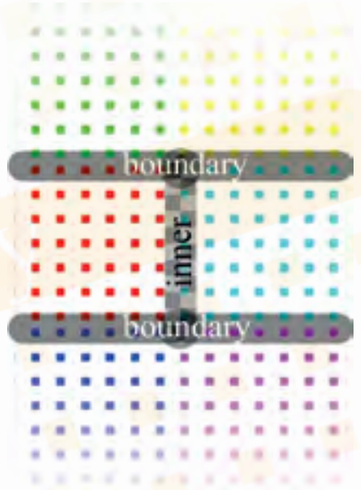


Figure 6: Inner and boundary elements in step 2.

new inner elements (cf. step 2 in Table 1 and Fig. 6) and they also can be eliminated in parallel on those four clusters. Combining these processors leads to a new boundary on which the next Schur complement can be computed. By repeating this procedure we end up with one processor set containing all processors. Thus we only have inner elements in the last step and the eliminating finishes the LU decomposition.

A second parallelization is realized by using all processors in each processor set. In step s there are 2^{s-1} processors in every cluster we can use to execute the local LU decomposition. In Fig. 7 we have an example for step 3, the distribution defines a block matrix,

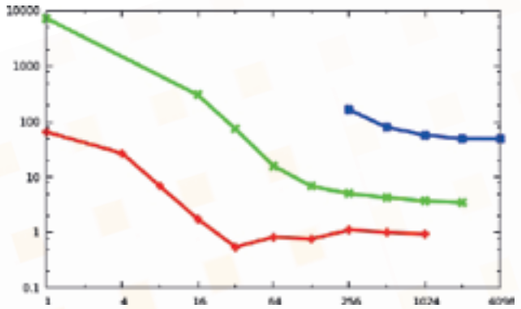
where all matrices are distributed block-columnwise to the processors. Here, in Z we have the inner elements, which should be eliminated, and in S are the boundary elements for the Schur

Figure 7: Distributed matrix on 4 processors in step 3.

complement. L and R define the connection between those elements. The Schur complement computed in this step is $S\text{-}LZ^{-1}R$, where a block LU decomposition of Z is created. Note, that there are no boundary elements in the last step. Thus, only a distributed Z matrix is given there.

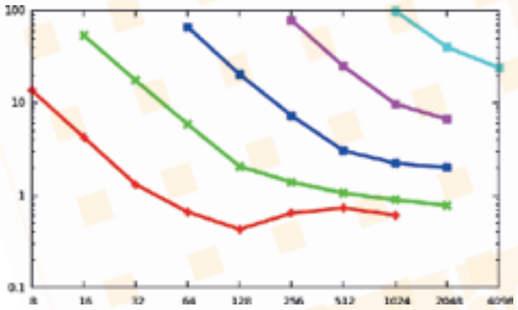
To execute the parallel LU decomposition, two basic communication routines are necessary. The first one is the one-to-one communication to fill the new ZRLS matrix from the two Schur complements in the step before. The second one is the broadcast routine to provide Z and L (after updating within the computation of the Schur complement) to all processors in the processor set.

Table 2: Parallel execution time on the CRAY XE6 Cluster in Stuttgart for the decomposition for the Poisson problem in the unit cube (3D).



	N = 35 937	N = 274 625	N = 2 146 689
P = 1	1:05.58 min.	2:03:25 hrs.	
P = 4	27.05 sec.		
P = 8	6.89 sec.		
P = 16	1.76 sec.	5:06.06 min.	
P = 32	0.55 sec.	1:13.94 min.	
P = 64	0.83 sec.	16.06 sec.	
P = 128	0.77 sec.	6.87 sec.	
P = 256	1.11 sec.	5.05 sec.	2:45.30 min.
P = 512	1.02 sec.	4.29 sec.	1:19.11 min.
P = 1024	0.93 sec.	3.73 sec.	49.75 sec.
P = 2048		3.44 sec.	49.75 sec.
P = 4096			49.55 sec.

Table 3: Parallel execution time on the CRAY XE6 Cluster in Stuttgart for the decomposition for the Stokes problem (2D).



	N = 80 131	N = 317 955	N = 1 266 961	N = 5 056 515	N = 20 205 571
P = 8	13.65 sec.				
P = 16	4.25 sec.	53.86 sec.			
P = 32	1.31 sec.	17.50 sec..			
P = 64	0.66 sec.	5.88 sec.	66.44 sec.		
P = 128	0.43 sec.	2.06 sec.	20.43 sec.		
P = 256	0.64 sec.	1.39 sec.	7.30 sec.	78.91 sec.	
P = 512	0.73 sec.	1.06 sec.	3.04 sec.	24.96 sec.	
P = 1024	0.61 sec.	0.89 sec.	2.23 sec.	9.76 sec.	98.87 sec.
P = 2048		0.78 sec.	1.99 sec.	6.63 sec	40.10 sec.
P = 4096					23.99 sec.

The Performance on Model Problems

We demonstrate the realization of the solver on two applications in 3D (Poisson) and 2D (Stokes), respectively. For the Poisson problem $-\Delta u = f$ in the unit cube $\Omega=(0,1)^3$ with homogeneous boundary conditions we use trilinear finite elements on a regular hexahedral mesh with mesh width $h_l=2^{-l}$ on refinement level $l \geq 0$. The mesh is distributed to $P=2^S$ processors by recursive coordinate bisection. We consider the Poisson problem up to refinement level $l=7$ and $N=2\,146\,689$ degrees of freedom and we compute the factorization for P=1–4096 processors, cf. Table 2. To decompose over 2 million unknowns in 3D, we need less than one minute on 1,024 processors.

A second realization in 2D is done for the Stokes problem $-\Delta u+\nabla p=0$, div $u=0$ on an L-shaped domain with inf-sup stable Taylor-Hood-Serendipity Q2/Q1 elements. This saddle point problem leads to a symmetric, but indefinite matrix A. Since this is a two dimensional problem with one dimensional interfaces, we obtain a high performance with the parallel solver, even for several thousand processors. The results are shown in Table 3. Over twenty million unknowns are decomposed in about 24 seconds on 4,096 processors.

For further information about the parallel direct solver, its programming scheme and more, see [1].

Acknowledgements

The authors acknowledge the financial support from BMBF grant 01IH08014A within the joint research project ASIL (Advanced Solvers Integrated Library) and the access to the Hermit cluster in Stuttgart.

References

[1] Maurer, D.
Ein hochskalierbarer paralleler direkter Löser für Finite Elemente Diskretisierungen, PhD thesis, Karlsruhe Institute of Technology, 2013.

[2] Wieners, C.
A geometric data structure for parallel finite elements and the application to multigrid methods with block smoothing, Comput. Vis. Sci., 13, pp. 161–175, 2010.

[3] Maurer, D., Wieners, C.
A parallel block LU decomposition method for distributed finite element matrices, Parallel Comput., 37, pp. 742–758, 2011.

contact: Daniel Maurer,
daniel.maurer@kit.edu

• Daniel Maurer

Karlsruher Institut für Technologie (KIT), Institut für Angewandte und Numerische Mathematik

Applications

Plasma Acceleration: from the Laboratory to Astrophysics

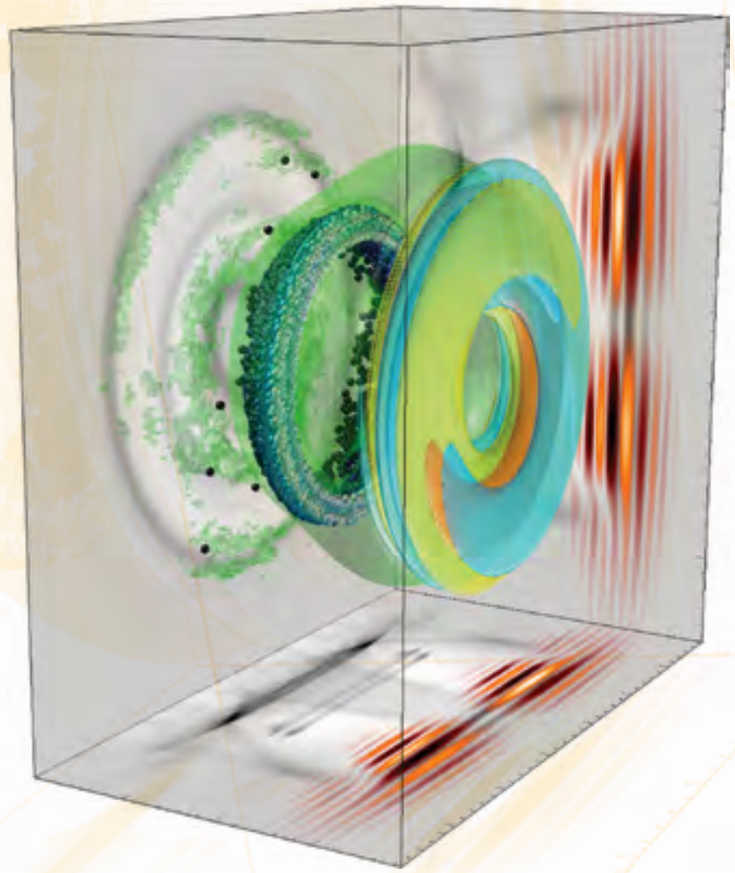


Figure 1: 3D simulation of a plasma wakefield excited by a doughnut shaped laser.

The study of novel particle acceleration and radiation generation mechanisms can be important to develop advanced technology for industrial and medical applications, but also to advance our understanding of fundamental scientific questions from sub-atomic to astronomical scales. For instance, particle accelerators are widely used in high-energy physics, and in the generation of x-rays for medical and scientific imaging. Although extremely reliable, conventional particle accelerator technology is based on radio-frequency fields, which can lead to very long and very expensive machines. For instance, the LHC at CERN is several tens of kilometers long and costed several billions of Euro. Thus, investigating new more compact accelerating technologies can be beneficial for science and applications. Plasmas are interesting for this purpose, because they support nearly arbitrarily large electric fields, and thus lead to a future generation of more compact accelerators. Plasma acceleration experiments, for instance, succeeded in doubling the energy of electron beams accelerated for several kilometers at SLAC in less than a one-meter long plasma.

Most of the matter in the known universe is also plasma. Hence, plasma based particle acceleration and radiation generation mechanisms may also play a key role in some of the astrophysical mysteries that remain unresolved, such as the origin of cosmic rays and gamma ray bursts through collisionless shocks and magnetic field generation and amplification mechanisms. Using computing resources at SuperMUC we have explored several questions with impact on the development of plasma acceleration technology and improving on our understanding of the mechanisms that may lead to particle acceleration and radiation in astrophysics. This report will describe some of the highlights that were achieved during our present allocation in these topics.

Plasma Acceleration in the Laboratory: towards Plasma-based Linear Colliders

As conventional particle acceleration techniques are hitting their technological limits, plasma based acceleration is emerging as a leading technology in future generations of higher energy, compact particle accelerators. Although initially proposed more than 30 years ago [1], the first ground breaking plasma acceleration experimental results appeared in 2005. Plasma acceleration is presently an active field of research, being pursuit by several leading laboratories (e.g. SLAC, DESY, RAL, LOA, LBNL). Electron or positron acceleration in plasma waves is similar to sea wave surfing. Plasma accelerators use an intense laser pulse or particle bunch (boats on water) as driver to excite relativistic plasma waves (sea waves for surfing). Accelerating structures are sustained by plasma electrons and are not affected by physical boundary effects as in conventional accelerators. The plasma accelerator only lasts for the driver transit time through the plasma. The resulting plasma wavelength is only a few microns long (<10 m for sea waves), and support accelerating electric fields up to 3 orders of magnitude higher than conventional particle accelerators. These accelerating fields can then accelerate electrons or positrons (as surfers in sea waves) to high energy in short distances (<1 m).

One of the challenges associated with the design of a plasma based linear collider is positron acceleration. Most important plasma based acceleration experimental results were performed in strongly nonlinear regimes, enabling to optimize the quality of accelerated electron bunches. It has been recognized, however, that these strongly non-linear regimes, are not suitable for positron acceleration. Using SuperMUC we then explored novel configurations for

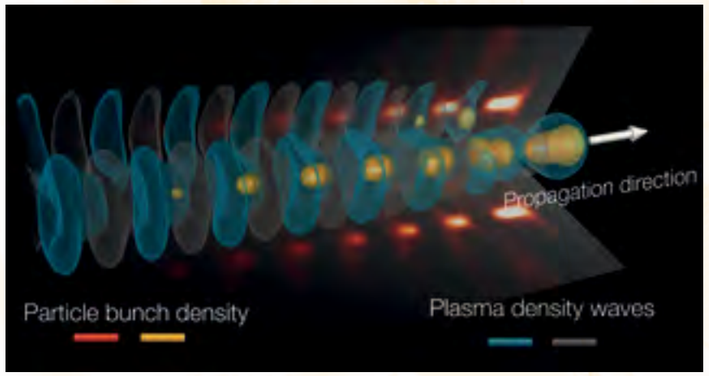


Figure 2: 3D simulation of a long particle bunch in a plasma leading to stable accelerating fields.

Applications

positron acceleration in strongly non-linear regimes. We investigated plasma acceleration driven by narrow particle bunch drivers and by doughnut shaped Laguerre-Gaussian lasers. Although the work using narrow particle bunch drivers is still in progress, we found that the wakefields driven by Laguerre-Gaussian beams can have good properties for positron acceleration in the strongly non-linear. Fig. 1 is simulation result showing a doughnut plasma wave. Typical simulations require 4×10^4 (doughout wakefields) - 2×10^5 (narrow drivers) core-hours, pushing 2.5×10^9 - 1.6×10^{11} particles for more than 4×10^4 time-steps.

We also performed simulations to clarify important physical mechanisms associated with a future plasma based acceleration experiment at CERN using proton bunches from the Super Proton Synchrotron (SPS) at the Large Hadron Collider (LHC) [2,3]. In this experiment, the proton bunch will drive plasma waves through a beam plasma instability called the Self-modulation Instability or SMI. The SMI modulates the bunch density profile radially. A fully self-modulated beam then consists of a train of shorter uniformly spaced bunches. Each of them will excite a plasma wave that grows through the beam, thereby producing acceleration gradients that grow through the beam. A long proton bunch will also be subject to the hosing instability or HI. The HI can lead to beam break up. Hence, HI suppression is required for successful experiments. SuperMUC simulations unraveled a new HI instability suppression mechanism, establishing the conditions for stable plasma wakefield generation in future experiments at CERN. Fig. 2 illustrates a fully self-modulated particle bunch propagating stably in the plasma, without HI growth. Typical simulations ran for roughly 5×10^4 core-hours, pushing 6.4×10^8 particles for more than 2×10^5 time-steps.

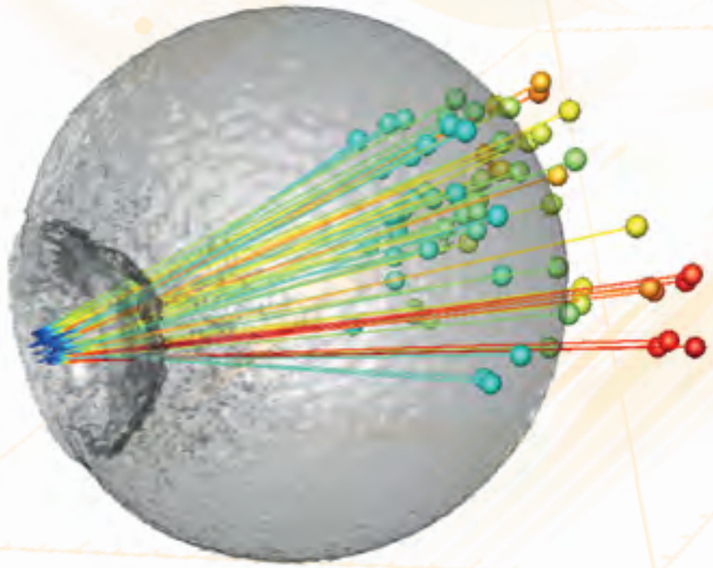


Figure 3: 3D simulation result showing the irradiation of a dense hydrogen target by a high intensity laser. Colored spheres represent accelerated protons.

Particle Acceleration in the Universe: mimicking Astrophysical Conditions in the Laboratory

The origin of cosmic rays and gamma ray bursts are fundamental mysteries for our understanding of the universe. These questions are closely related to particle acceleration in shocks and to strong magnetic field generation. Collisionless shocks have been studied since decades in the context of space and astrophysics due to their potential of efficient particle acceleration to energies larger than 10^{15} eV. Large-scale magnetic fields are also important for particle acceleration in astrophysics, but also allow for non-thermal radiation processes to operate. Exploring these extreme scenarios is complex, since astronomical observations are limited.

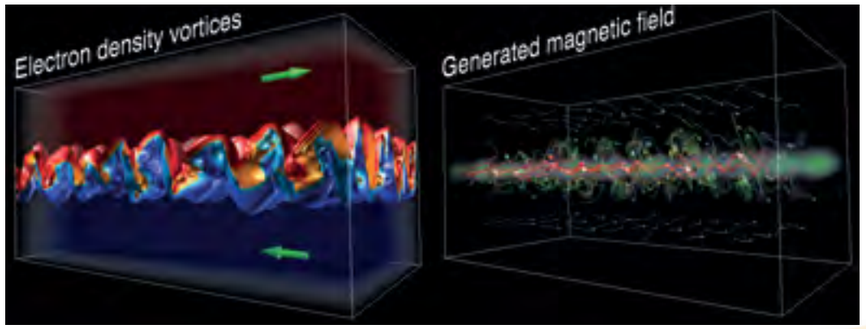


Figure 4: 3D simulation result showing electron density vortices and magnetic field due to the KHI.

Recent experiments, however, have started to investigate the onset of collisionless shocks in the laboratory resorting to high power laser pulses. There are distinct shock types according to the energy transfer to the plasma. We performed simulations in SuperMUC that enabled to explore the transition between different shock types. Our simulations showed that electrostatic shocks, which lead to the formation of strong electric fields, can be used to efficiently accelerate plasma ions or protons (see Fig. 3). Electromagnetic shocks, which lead to the formation of intense magnetic fields, can lead to particle acceleration in astrophysics through Fermi-like acceleration mechanisms. Magnetic field generation also occurs in electromagnetic shocks, due to the Weibel or Current Filamentation Instability (WI or CFI). These instabilities can grow when counter-streaming plasma flows interpenetrate. Magnetic field generation and amplification is important in astrophysics as they can lead to strong bursts of radiation. Thus, in addition to magnetic field amplification in collisionless shocks, simulations at SuperMUC enabled the discovery of a novel mechanism driving large-scale magnetic fields in shearing counter-streaming plasma flows. Velocity shear flows lead to the development of the Kelvin-Helmholtz Instability or KHI. We found that the KHI is an important dissipation mechanism, capable to efficiently transform plasma kinetic energy into electric and magnetic field energy [4]. Fig. 4 shows the development of the characteristic KHI vortices and associated magnetic field in conditions relevant for astrophysics. Each simulation took 2×10^4 core-hours pushing 10^{10} particles for more than 10^6 iterations.

On-going Research / Outlook

SuperMUC enabled to make important advances to plasma acceleration in the laboratory and in astrophysics in conditions for which purely analytical models are currently unavailable. We managed to perform very large simulations, which would not have been possible in smaller supercomputers.

References and Links

[1] **Tajima, T. and Dawson, J.**
Phys. Rev. Lett. 43 267 (1979).

[2] **Caldwell, A. et al.**
Nat. Physics 5, 363 (2009).

[3] **http://awake.web.cern.ch/awake/**

[4] **Grismayer, T. et al.**
Phys. Rev. Lett. 111 015005 (2013).

contact: Jorge Vieira,
jorge.vieira@ist.utl.pt

- Jorge Vieira
- L.D. Amorim
- Paulo Alves
- Anne Stockem
- Thomas Grismayer
- Luís Silva

GoLP/IPFN,
Instituto Superior Técnico, Lisbon, Portugal

Towards Simulating Plasma Turbulences in Future Large-Scale Tokamaks

GEM is a 3D MPI-parallelised gyrofluid code used in theoretical plasma physics at the Max Planck Institute for Plasma Physics, IPP, at Garching near Munich, Germany. Recently, IPP and LRZ collaborated within a PRACE Preparatory Access Project to analyse various versions of the code on the HPC systems SuperMUC at LRZ and JUQUEEN at Jülich Supercomputing Centre (JSC) to improve the weak scalability of the application [1].

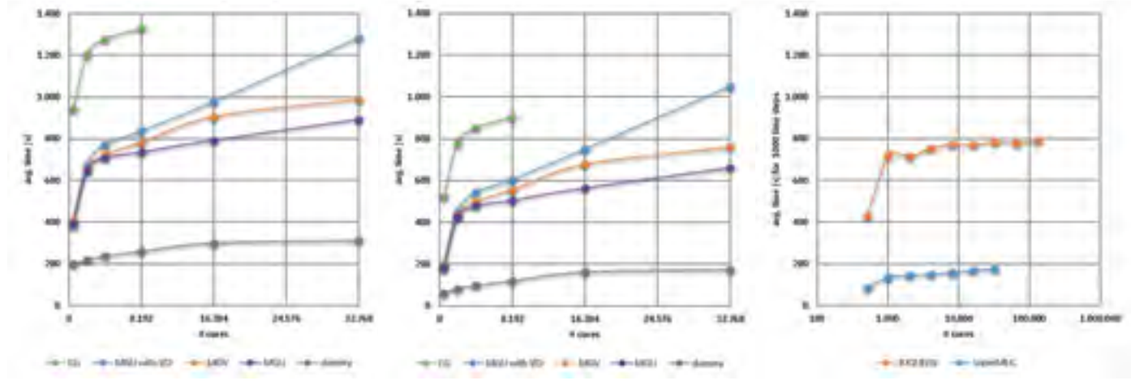
Simulation of Turbulence in Plasmas

The code GEM addresses electromagnetic turbulence in tokamak plasmas [2]. Its main focus is on the edge layer in which several poorly understood phenomena are observed in experiments. The code is written in Fortran/MPI and is based on the electromagnetic gyrofluid model. It can be used with different geometries depending on the targeted use case. The code has been run up to now on conventional tokamak cases like the ASDEX-Upgrade [3] at the IPP in Garching with a plasma minor radius r = 0.5 m (see Fig. 1). A typical grid to cover the edge layer for conventional tokamaks is 64 × 4,096 × 16. To simulate thin-strip, medium-to-kamaks like the JET tokamak (r = 1.25 m) in Culham, UK [4], or a large-scale tokamak like the ITER reactor-plasma experiment (r ≈ 2 m) currently under construction in Cadarache, France [5], grid sizes up to 1,024 × 16,384 × 64 are necessary.

Figure 1: Plasma vessel of the ASDEX-Upgrade tokamak at IPP.
© IPP, Photo: Volker Rohde

Scaling Towards Large-Tokamak Cases

The main goal of the project was to establish weak scalability of the gyrofluid code GEM to larger systems: if the thin-strip, small-tokamak case 64 × 4,096 × 16 can be afforded on 512 cores, the largest planned thick-strip, large-tokamak case 1,024 × 16,384 × 64 should be able to run on 131k cores in the same or similar wall clock time. Since the main bottleneck of the code is in the solver, we have particularly focused on its improvement. Various versions of the MPI-parallelised code have been set up for the weak scaling analysis. In most cases the I/O routines were eliminated to solely concentrate on the solver features. The dummy version functioning as a baseline for our measurements uses no boundary or sum information. The CG version implements a conjugate-gradient method for the solver iteration. Furthermore, two different Multigrid versions have been analysed: the MGV version (Multigrid with V-cycle scheme) and the MGU version (Multigrid with U-cycle scheme) which has been developed within this project.



(a) (b) (c)

Figure 2: (a) Overall wall-clock time for the various versions of the code on SuperMUC. (b) Time spent in the MPI functions (from the Scalasca analysis). (c) Comparison of the weak scaling of the MGU version of GEM on SuperMUC and on JUQUEEN.

A comparison of the performance of the various versions on SuperMUC is shown in Fig. 2 (a). Detailed performance analysis using the Scalasca tool [6] revealed that the scalability of the dominating MPI routines determines the scalability of the entire code, as shown in Fig. 2 (b). This is very significant to the CG solver due to its heavy usage of the MPI_Allreduce function. Here the MGU version shows best scalability. Finally, Fig. 2 (c) presents the good weak scaling behaviour of the MGU version (excluding the I/O) up to 32,768 cores on SuperMUC (IBM System x iDataPlex) at LRZ and up to 131,072 cores on JUQUEEN (IBM Blue Gene/Q) at Jülich Supercomputing Centre.

The code has been run up to now on conventional tokamak cases like the ASDEX-Upgrade [3] (see Fig. 1) at the IPP in Garching with a plasma volume V = 13 m^3. A typical grid to cover the edge layer for conventional tokamaks is 64 × 4,096 × 16. To simulate thin-strip, medium-tokamaks like the JET tokamak (V = 155 m^3) in Culham, UK [4], or a large-scale tokamak like the ITER reactor-plasma experiment (V ≈ 840 m^3) currently under construction in Cadarache, France [5], grid sizes up to 1,024 × 16,384 × 64 are necessary.

Acknowledgements

Our work was financially supported by the PRACE project funded in part by the EUs 7th Framework Programme (FP7/2007-2013) under grant agreement nos. RI-283493 and RI-312763. The results were obtained within the PRACE Preparatory Access Type C Project 2010PA1505 “Scalability of gyrofluid components within a multi-scale framework”.

References

[1] **Scott, B., Weinberg, V., Hoenen, O., Karmakar, A., Fazendeiro, L.**
Scalability of the Plasma Physics Code GEM, PRACE Whitepaper, http://www.prace-project.eu/IMG/pdf/wp125.pdf.

[2] **Scott, B., et al.**
Phys. Plasmas 12 (2005) 102307, Plasma Phys. Contr. Fusion 48 (2006) B277, Plasma Phys. Contr. Fusion 49 (2007) S25, Contrib. Plasma Phys. 50 (2010) 228, IEEE Trans Plasma Sci. 38 (2010) 2159.

[3] http://www.ipp.mpg.de/16195/asdex

[4] http://www.ccfe.ac.uk/JET.aspx

[5] http://www.iter.org

[6] http://www.scalasca.org

contact: Volker Weinberg, volker.weinberg@lrz.de

- Volker Weinberg
- Anupam Karmakar

Leibniz Supercomputing Centre (LRZ)

A Flexible, Fault-tolerant Approach for Managing Large Numbers of Independent Tasks on SuperMUC

In many HPC applications, small computational tasks have to be re-iterated a thousand times or more. For example, to create a computer animated movie, several ten thousand single frames have to be rendered. Likewise, in genetic research, millions of gene sequences need to be aligned. A common aspect is that the applied software usually only scales up to one compute node (16 cores on SuperMUC). The runtime of the individual tasks is short and researchers often resort to creating batch jobs for each task and submitting them to a scheduler like SLURM or SGE. This is a convenient and efficient solution if the number of users on the system is small. On HPC resources like SuperMUC with thousands of users, however, this procedure is not feasible since the number of jobs per user is limited (on SuperMUC, max. 3 jobs can run simultaneously per user). The standard way around this limit are job farming approaches.

In this article we describe a novel approach to job farming. It is based on an independent data base run by the user. This makes it much more flexible and easier to handle than most job farming tools integrated into batch systems whose configuration cannot be changed by the user. As a by-product, it features several extremely convenient properties like fault-tolerance, dynamic resizing of the attached resources, and a mechanism for backfilling which makes it preferable over conventional job farming approaches.

Orchestrating Thousands of Worker Nodes using Redis–a No-SQL Data Base

The concept of our job farming approach is based on the light-weight data base server redis. This no-SQL data base can handle large numbers of queries by thousands of different users efficiently. In our example this redis server will run on a SuperMUC login node which can be accessed by the compute nodes as well. The data base server will store and manage the subtasks that will be processed by the compute nodes.

In our illustration example we use the approach for rendering a movie. Table 1 was created at the concept stage of the movie. It contains columns like 'Scene Description' which are not needed as input to render the animation. The important information are the file names as well as the variables 'Start Frame' and 'End Frame'. Together they describe several thousand input parameters and input data for the rendering software Blender (www.blender.org). Rendering one individual frame is done using a python script (e.g. rendering frame no. 22 of the first scene using 32 threads):

```
lrz_render_frame.py –l 1 –f 22 –t 32
```

Based on the command line parameters and the description in Table 1, this python file creates the correct system call to render this specific frame in Blender. Running the python script from Fig. 1, one can create similar command strings for each frame that has to be rendered. These strings are sent to the redis server ('redis-cli lpush'). Eventually, the redis data base stores several thousand string keys representing the command lines for all frames.

Filename	Start Frame	End Frame	Scene Description
01_01_opening.blend	1	1400	Flight over Campus
02_01_node.blend	1	415	One Node of SuperMUC
02_02_cpu.blend	1	240	Zoom into a CPU Core

Table 1: Excerpt from the task description file for a computer animated movie.

Now, one may spawn an almost arbitrarily large number of worker processes on SuperMUC in order to perform the rendering of the frames.

The only requirement for the worker processes is a connection to the redis server and their number is only restricted by the total number of frames, i.e. one could use all of SuperMUC at once to render the movie using all 155,656 cores. The simplest way to start the worker processes is a LoadLeveler script invoking poe with the shell script 'startworker.sh' in Fig. 2 as argument. This script invokes a single process which iteratively fetches a single string from the redis server and executes it as a shell command.

Since the software Blender has an integrated OpenMP parallelization, it will efficiently use the 16 physical cores (or 32 logical cores using hyperthreading) available in one compute node of SuperMUC.

```
## the 2-d array filmdata contains the content of Table 1
##
for line_number in range(1,len(filmdata)):
    mydata          = filmdata[line_number]
    start_frame     = mydata[1]
    end_frame       = mydata[2]
    for frame_number in range(int(start_frame),int(end_frame)+1):
    os.popen('redis-cli lpush job_queue_1 "lrz_render_frame.py
    -l %s -f %s -t 32"'%(line_Number, frame_number))
```

Figure 1: Python script for sending command strings for each frame to the redis server.

```
#!/bin/bash
# host: localhst port: 6379
. /etc/profile
module load redis
while :
do
$(echo $(redis-cli -h $1 -p $2 brpop job1 0) | cut -d' ' -f2-)
done
```

Figure 2: Bash script 'startworker.sh' for a redis worker process which fetches command strings from the redis server and executes them in an endless loop.

The proposed work flow is probably one of the simplest implementations for a hybrid MPI/OpenMP programming model. Once Blender has finished rendering a frame, its parent process will reconnect to the redis server and fetch a new command string. This means, that the described work flow has an integrated load balancing mechanism.

Since one may spawn thousands of these worker processes , one can easily render a complete movie overnight.

doRedis–Embedding the Redis Workflow in Conventional Loop Syntax

In its pure version, the redis workflow might be considered cumbersome because several scripts have to be run on different machines and the communication with the redis server has to be performed 'manually'. Within the R programming language (www.r-language.org), an easy-to-use package for the redis workflow has been developed, doRedis. Apart from requiring a redis server running in the background, the workflow is completely embedded in the conventional R syntax. As in the official example for doRedis, we will illustrate the usage by a simple example in which π is approximated.

Consider a circle of radius 1 in Fig. 4 which circumscribes an area equal to π. Now, we draw two random numbers from the interval [-1,1], representing arbitrary x- and y-coordinates. The probability P that the new point is located within the circle equals $\frac{\pi}{4}$ (the area of the circle divided by the area of the surrounding square of length 2). We can approximate this probability P by drawing a large number of new points (x,y) and checking whether they are inside the circle, i.e. $\sqrt{x^2+y^2}<1$.

In R, we can implement this using the code example in Fig. 5, which are based on a simple function for drawing 10 million observations, sample_10M. Additionally, we have a loop statement repeating the function B times. In the redis version, these B iterations are sent to the local worker processes via the foreach statement of doRedis.

Apart from lines 1 to 3 which set up the connection to the redis server both versions are almost identical. Line 3 starts local R worker processes. Similar to the previous example, these processes fetch strings from the data base and automatically execute these strings as normal R-code. However, the mechanism in doRedis executes these strings in a more elaborated way. For example, it has an integrated implementation for fault tolerance. If a worker does not return a result within a certain time, the subtask will be automatically rescheduled assuming that a failure has occurred on the original worker node responsible for the subtask.

Figure 3: Redis-based workflow for rendering a movie on SuperMUC using the software Blender.

Once lines 1 to 3 have been run, the redis environment for parallel processing of tasks is established. Subsequently, one may run further parallel computations like the one in lines 5 to 17. It is also worth noting that one can transform the redis version in (b) into a serial version simply by replacing %dopar% by %do%. In this case, R will process the foreach statement as a normal loop statement which is a convenient feature if the worker processes are temporarily not available.

Another Use Case: Gene Sequence Alignment

In gene sequence analysis, we have to align millions of sequences to a reference genome. A well-established software for this problem is Blast (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Similar to Blender it provides a good shared

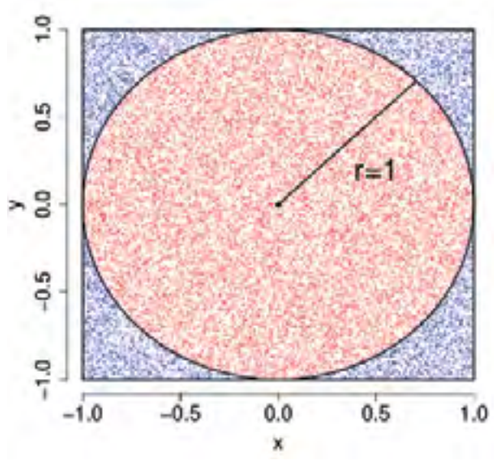


Figure 4: Approximating by Monte Carlo Integration

memory parallelization, but an MPI implementation is missing and not expected to yield strong benefits. It is far more efficient to distribute the millions of alignments over many computing nodes. In comparison to the previous case, the alignment features a data base which has to be copied to each node at the beginning of the computation. Fig. 6 shows a Gantt chart for 2,048 Blast processes. In this case, several Blast instances were run on the SuperMUC compute nodes. The green region at the beginning represents the copying of the reference

(a) serial version

```
1
2
3
4
5  sample_10M <- function() {
6      x <- runif(1e+07, min = -1, max = 1)
7      y <- runif(1e+07, min = -1, max = 1)
8      mean(sqrt(x^2 + y^2) <= 1)
9  }
10
11 B <- 100
12 portion_inside <- numeric(B)
13
14 for (i in 1:B)
15 portion_inside[i] <- sample_10M()
16
17 (pi <- 4 * mean(portion_inside))
```

(b) redis version

```
1  library(doRedis)
2  registerDoRedis(queue = "pi3", host = 'localhost')
3  startLocalWorkers(n = 4, queue = "pi3")
4
5  sample_10M <- function(j) {
6      x <- runif(1e+07, min = -1, max = 1)
7      y <- runif(1e+07, min = -1, max = 1)
8      mean(sqrt(x^2 + y^2) <= 1)
9  }
10
11 B <- 100
12
13
14 portion_inside <-
15 foreach(j = 1:B) %dopar% sample_10M(j)
16
17 (pi <- 4 * mean(unlist(portion_inside)))
```

Figure 5: Code examples in the R computing language to approximate. (a) serial version and (b) redis code.

genome data base to the main memory of each participating SuperMUC node. The data base was copied to RAM-Disk (/dev/shm), i.e. it was used as an in-memory data base. Several Blast processes were run on each node, and these processes could share the in memory-data base. This strategy significantly reduces the amount of RAM per core although it moderately slows down the computation in comparison to the serial version. The panel (a) in Fig. 6 corresponds to a serial version which performed ten alignments (iterating orange and blue regions).

The panel (b) in Fig. 6 depicts the parallel version using the proposed redis workflow. The green region for copying the data shows that certain nodes receive the data base earlier and directly start their alignment tasks. Subsequently, the ten alignments (orange and blue) on the compute nodes take slightly longer than in the serial version. Nonetheless, using this strategy 20,000 alignments could be performed in roughly the same time in which the serial version only managed to perform 10 alignments.

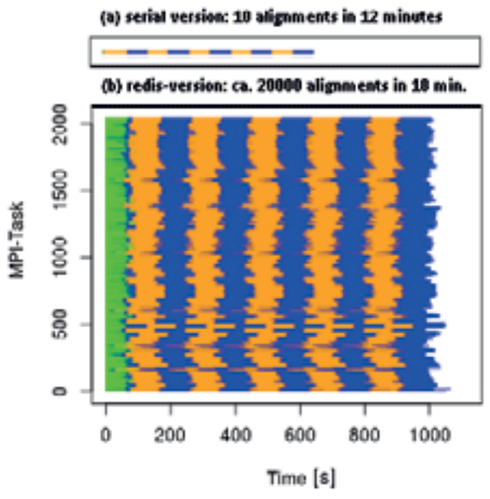


Figure 6: Gantt chart for gene sequence alignment using Blast (a) serially and (b) the redis-based parallel version.

Discussion and Outlook

In comparison to most other job farming approaches, the redis work flow described here offers several desirable features like flexibility and extensibility.

Typical job farming approaches that are integrated into standard batch systems have no automatic load balancing. In the redis work flow, each worker immediately queries a new task as soon as it has finished processing its previous task. Thus, if tasks with varying runtime exist, compute nodes that finished a task will automatically perform another task. Likewise, if we have tasks with different requirements (e.g. main memory), one can define several job queues on the redis server which will be processed by different workers. The R-connector doRedis also provides a fault-tolerant scheduling mechanism. It will automatically restart a task on another worker node if it either loses the connection to one of the workers or an optional, user-defined timeout occurs.

Last but not least, the attachment of the workers as well as the submission of tasks are completely dynamic. It is even possible to start workers before the data base is filled with tasks. The worker processes will simply remain idle until tasks are sent to the redis server. Additionally, it is possible to add tasks during runtime–one simply has to send additional tasks to the job queue on the redis server. This provides a convenient way for implementing 'backfilling', i.e. assigning new work to resources that currently idle because they are waiting for other processes of the same job to finish. Similarly, one may assign additional resources to a job queue to speed up the work. This feature also enables cloud bursting.

Figure 7: Implementation of a hybrid cloud using the redis approach.

For example, a user may outsource the computations to the resources at the LRZ when their requirements exceed the capacities of the local resources at the institute of the user (e.g. during peak hours). During off-peak hours, the computations can be performed on designated local hardware at the institute of the user.

Another interesting extension of the redis work flow is worth mentioning. The presented redis approach is capable of combining very heterogeneous resources with various kinds of architectures. The authors have already conducted several test runs where resources of different University departments, the commercial cloud services Amazon EC2, and LRZ systems jointly performed data mining, where several thousand machine learning algorithms had to be run. The tasks were distributed among the different resources by a redis server running on a local PC to which all compute nodes had access. The machine learning algorithms were implemented in R. This allowed us to include diverse hardware architectures into the work flow, establishing a fault-tolerant, dynamic, hybrid cloud.

contact: Christoph Bernau,
christoph.bernau@lrz.de

- Christoph Bernau
- Ferdinand Jamitzky
- Helmut Satzger

Leibniz Supercomputing Centre (LRZ)

SHAKE-IT: Evaluation of Seismic Shaking in Northern Italy

Ground shaking due to an earthquake not only depends on the energy radiated at the source but also on propagation effects and amplification due to the response of geological structures. A further step in the assessment of seismic hazard, beyond the evaluation of the earthquake generation potential, requires then a detailed knowledge of the local Earth structure and of its effects on the seismic wave field. The simulation of seismic wave propagation in heterogeneous crustal structures is therefore an important tool for the evaluation of earthquake-generated ground shaking in specific regions, and for estimates of seismic hazard.

Current-generation numerical codes, and powerful HPC infrastructures, now allow for realistic simulations in complex 3D geologic structures. We apply such methodology to the Po Plain in Northern Italy, a region with relatively rare earthquakes but with a large property and industrial exposure – as it became clear during the recent events of May 20-29, 2012. The region is characterized by a deep sedimentary basin: the 3D description of the spatial extent and structure of sedimentary layers is very important, because they are responsible for significant local effects that may substantially amplify the amplitude of ground motion. Our goal has been to produce estimates of expected ground shaking in Northern Italy through detailed deterministic simulations of ground motion due to possible earthquakes.

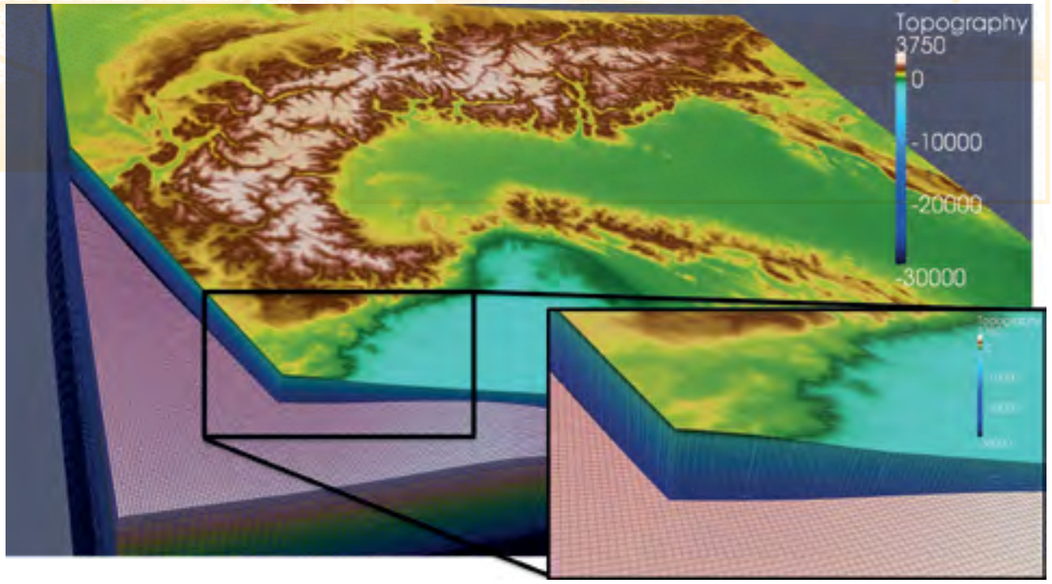


Figure 1: Mesh of the crust and the mantle, coloured as a function of topography. The mesh honours the topography of free surface and crust/mantle boundary.

Results

We started the work with the implementation of a 3D description of the subsurface geological units of the sedimentary Po Plain basin, plus neighbouring Alps and Northern Apennines mountain chains. Once the spatial characters of the structure have been set by merging detailed information from scientific literature and databases, the elastic parameters and density of the different units had to be adjusted. Our strategy consisted of defining an a priori plausible space of model parameters, and exploring it following a metaheuristic procedure based on a Latin hypercube sampling. This strategy was chosen to direct the trial-and-error procedure, as it provides a more uniform coverage of the subspace of interest than a random sampling. Each model has been tested for its ability to reproduce observed seismic waveforms for recent events. Best models were those providing the best fit. We have extensively simulated and independently compared in performance three different models for the Earth's crust – ranging from a simple 1D model, to a 3D large-scale reference model, to our new "optimal" high resolution 3D model – to gauge their ability to fit observed seismic waveforms for a set of 20 recent earthquakes occurred in the region. We observe a strong improvement of the fit to observed data by the wave field computed in our new detailed 3D model of the plain. Specifically, the new model has shown to be able to reproduce the long so-called 'coda' of the signals – late-arriving energy, due to reverberations and scattering. A 1D model, as expected, cannot account for the waveform complexity due to the effect of the sediments, while the high-resolution 3D model does a very good job in fitting the envelope of the seismograms. For the local earthquakes for which high-quality seismograms were available, the new 3D model can reproduce well the peak ground acceleration at periods longer than 2 s, and the overall duration of the shaking – two parameters of highest relevance for engineering purposes, as they are of great consequence to model the response of high-rise buildings and soil liquefaction effects. The new high-resolution model shows therefore significant improvement with respect to both the 1D and the large-scale 3D models. Understandably, the differences are most evident for propagation paths crossing the sedimentary basin.

Having verified that we now have a model (and a reliable computational framework), able to reproduce the gross characters of seismically-induced ground motion, we can then put them to work to infer the ground shaking that would be caused by seismic sources deemed plausible for the region. Seismic hazard assessment is built on the knowledge of past activity derived from historical catalogues and geological evidence. We can legitimately reason that some event of the past will repeat itself with very similar characters in the future, and the numerical framework (that we have verified on recent events) permits to reconstruct the ground shaking that will ensue. We have therefore considered in detail two specific cases. One refers to the earthquake that struck the city of

Ferrara in 1570, an event very similar to the one occurred in May, 2012; the other to a series of historical events, with similar characters, located along the same linear geological structure roughly comprised between the cities of Modena and Parma. The uncertainty in the knowledge of source parameters – quite relevant in the case of knowledge deriving from historical catalogs – has been explicitly considered by generating suites of source parameters spanning the uncertainty ranges of hypocentral coordinates and geometrical parameters. For each earthquake, we represent the uncertainty regions by 200 instances generated through Latin hypercube sampling, and analyse the statistical properties of the resulting ensemble of wave fields generated. We may thus address the minimum, maximum, or median of the suite of peak ground accelerations, for all the geologically acceptable source models, in all possible locations. We can also explore the dependence of shaking parameters on source parameters – such as hypocentral depth or fault plane orientation. Different geological settings (i.e. hard rock, vs. consolidated sediments, or water-saturated sediments) in fact show a quite different response.

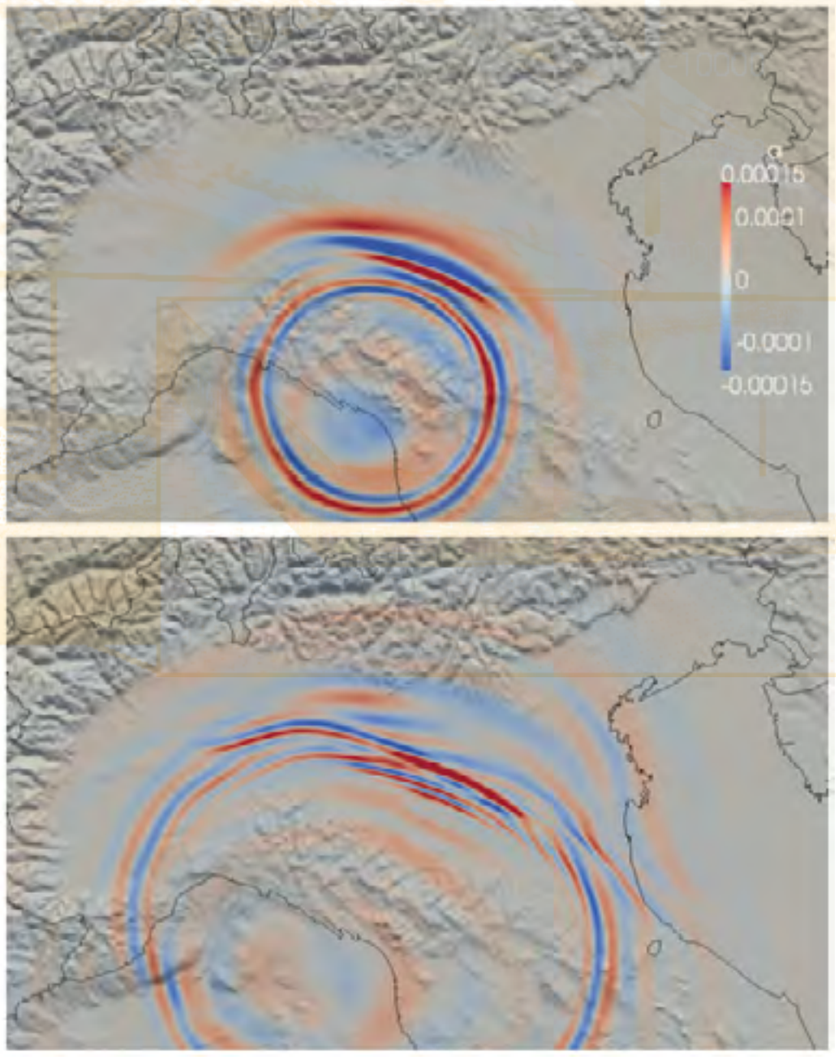


Figure 2: Snapshots at 30s and 70s of the simulation of the Mw=5.2 June 21, 2013, earthquake. The effect of the basin structure is clearly apparent when the wave front passes the ESE-WNW-trending boundary with the mountain front.

On-going Research / Outlook

This study builds the basis for a physics-based approach to seismic hazard estimates in Northern Italy. We will need, on one side, to increase the resolution of the geological model, in some critical locations, to slightly increase the highest significant frequency. We are currently limited to relatively low-frequency ($f < 0.5$ Hz) and more efforts are needed to go to higher frequency. A more detailed description of the geological structure may permit to increase to $f \sim 1$ Hz. Stochastic synthesis is required to reach further high frequencies – $f > 1$ Hz, with engineering interest for low-rise residential buildings – so a hybrid deterministic-stochastic approach must be used.

References and Links

[1] **Molinari, I., Morelli, A.**
Geophys. J. Int., 185, 352-364, doi: 10.1111/j.1365-246X.2011.04940.x, 2011.

[2] **Morelli, A., Molinari, I., Basini, P., Danecek, P.**
Abstract S32B-02 presented at 2012 Fall Meeting, AGU, San Francisco, Calif. (USA), 3-7 Dec., 2012.

[3] **Molinari, I., Morelli, A., Basini, P., Berbellini, A.**
Abstract S32B-03 presented at 2013 Fall Meeting, AGU, San Francisco, Calif. (USA), 9-13 Dec., 2013.

[4] **Tondi, R., Cavazzoni, C., Danecek, P., Morelli, A.**
Computers & Geosciences, 48, 143-156, DOI: 10.1016/j.cageo.2012.05.026, 2012.

[5] **Gualtieri, L., Serretti, P., Morelli, A.**
Geochemistry, Geophysics, Geosystems, DOI:10.1002/2013GC004988, 2013.

contact: Andrea Morelli,
andrea.morelli@bo.ingv.it

- Andrea Morelli [1]
- Peter Danecek [1]
- Irene Molinari [1]
- Andrea Berbellini [1]
- Paride Lagovini [1]
- Piero Basini [2]

[1] Istituto Nazionale di Geofisica e Vulcanologia, Italy

[2] Toronto University, Toronto, Canada

Quantum Photophysics and Photochemistry of Biosystems

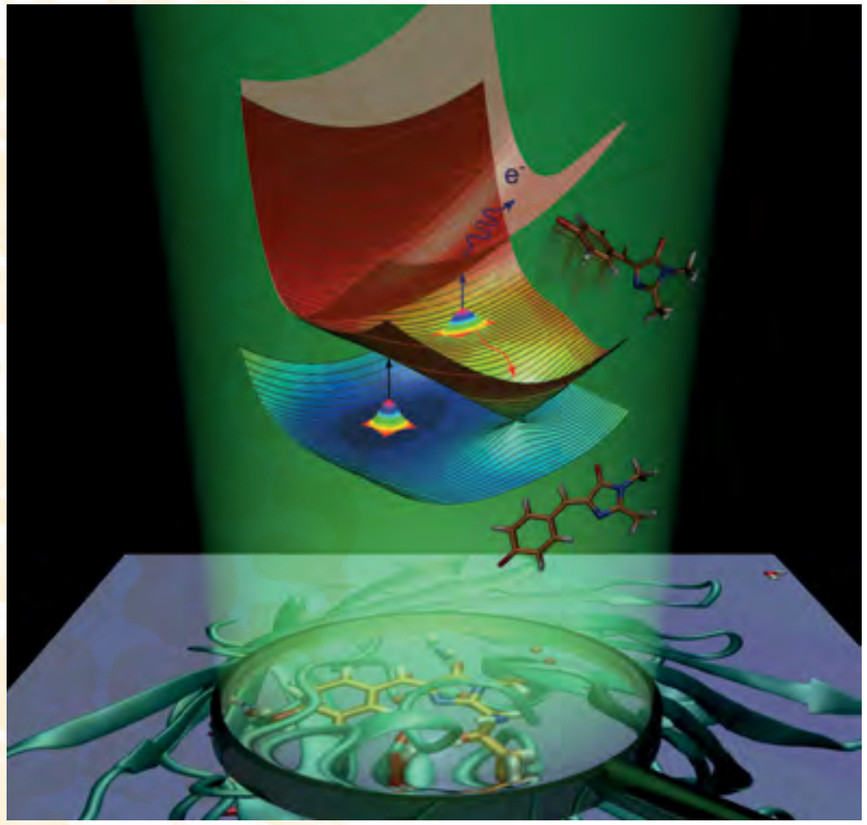


Figure 1: When taken out of the Green Fluorescent Protein, its anionic light-absorbing molecular unit shows a remarkable fundamental interplay between electronic and nuclear excited-state decay channels in the ultrafast photo-initiated dynamics occurring on a pile of potential energy surfaces.

The interaction of molecules with light is central to vital activity of living organisms and human beings. Photosynthesis, vision in vertebrates, solar energy harvesting and conversion, and light sensing are remarkably efficient processes, and much focus has been on elucidating the role played by the protein environment in their primary events, which occur on a timescale down to sub-picoseconds. Our project deals with computationally demanding simulations of excited-state evolution of photo-active proteins and their light-absorbing molecular units using highly correlated multi-reference methods of Quantum Chemistry (QC) and their efficient parallel implementation. Such large-scale calculations combined with a high accuracy have been enabled through the PRACE LRZ's SuperMUC infrastructure and the highly tuned and optimized design of the Firefly package [1].

Results

High-level QC methods are required for describing photo-initiated quantum molecular dynamics that involves multiple electronic states. It is challenging to describe electronic fast and efficient de-excitation occurring through so-called conical intersections, where topography around an intersection seam of two degenerate electronic states influences the transition between them. Systems with quasi-degeneracy require multi-reference approaches. The new XMCQDPT2 method [2] is a unique and invariant approach to multi-state multi-reference perturbation theories (PT), allowing an accurate description of large and complex systems; in particular, near the points of avoided crossings and conical intersections. The use of high-performance high-core-count clusters becomes mandatory requirement for predictive modeling of such systems.

Unlike well-established approaches in classical molecular dynamics, QC methods deserve special attention when implementing them on modern computer architectures. Huge data sets of 4-indexed two-electron (2-e) integrals are generated and stored on a disk after their evaluation and intermediate sorting in a two-pass transformation. These data are subsequently used in computationally intensive parts of the code, where most of the computational efforts are due to summation of the individual terms of the PT series. Thus, the QC algorithms inevitably contain different stages, which are I/O intensive (e.g., integral transformation) and computationally intensive (e.g., direct summation of the PT series). The I/O performance is, therefore, of utmost importance. Besides, a theoretical/formulae complexity requires a careful tuning of computationally intensive parts of the code. A straightforward implementation of the summation of the PT series is very inefficient, since at least one or more slow divide operations are required to calculate each individual term of the PT series. Moreover, the summation runs over a large amount of data, involving some combinations of transformed 2-e integrals, thus making such algorithm not processor cache-friendly.

The efficient implementation of the XMCQDPT2 theory within the Firefly package is based on a developed family of the cache-friendly algorithms for the direct summation of the PT series [3], as well as on a so-called resolvent-fitting approach [4], which uses a table-driven interpolation for the resolvent operator, thus further reducing computational costs.

Being executed in its recently developed extreme parallel (XP) mode of operations, Firefly is scalable up to thousands of cores. The XP mode includes a three-level parallelization (MPI process → groups of weakly coupled MPI communicators with parallelization over MPI inside each group → optional multithreading within each instance of a parallel process) and efficient asynchronous disk I/O with real-time data compression/decompression (throughput of up to 5-7 GByte/s).

The Firefly package enables to perform efficient large-scale XMCQDPT2 calculations for systems with up to 4,000 basis functions and with large active spaces, which comprise several millions of configuration state functions. Protein and solution environments may also be taken into account using multi-scale approaches. The overall size of model systems may comprise 5,000–10,000 atoms.

The project has the total allocated CPU time of 3,885,000 core-hours, which have entirely been used. The average number of cores used by a single run is 1,024 and up to 2,944 with at least 24 hours duration. A distributed high-capacity high-bandwidth file system with up to 100 TBytes for temporary storage during a single run is required. The IBM GPFS file system of the LRZ's SuperMUC has been found to be the most sustained to a high load caused by the most I/O intensive parts of the XMCQDPT2 code. The direct benefits from tuning and optimizing the Firefly code have been reaped during the project. The most I/O intensive parts of the code have been redesigned targeting particular large XMCQDPT2 jobs.

The following scientific and technical milestones have been achieved within the project:

- By using excited-state electronic structure calculations of the isolated anionic Green Fluorescent Protein (GFP) chromophore, we have been able to disclose a striking interplay between the electronic and nuclear dynamics that is coupled with a remarkable efficiency in the ultrafast excited-state decay channels of this chromophore (Fig. 1). The energy exchange is found to be fast and, importantly, mode-specific [5]. This finding lays the groundwork for mode-selective photochemistry in biosystems.

- By simulating the photo-initiated early-time nuclear dynamics of the GFP and KFP proteins, their absorption spectral profiles have been retrieved, which appear to be remarkably similar to those of the isolated chromophores [5]. Here, high-frequency stretching modes play an important role defining the spectral widths. Remarkably, they also serve as reaction coordinates in photo-induced electron transfer, which competes with internal conversion mediated by twisting. Many GFP-related proteins prohibit internal conversion, thus possibly directing the decay towards electron transfer; whereas fluorescence may be regarded as a side channel only enabled in the absence of relevant electron acceptors. This finding adds to our understanding of biological evolution and function of light-sensitive proteins.

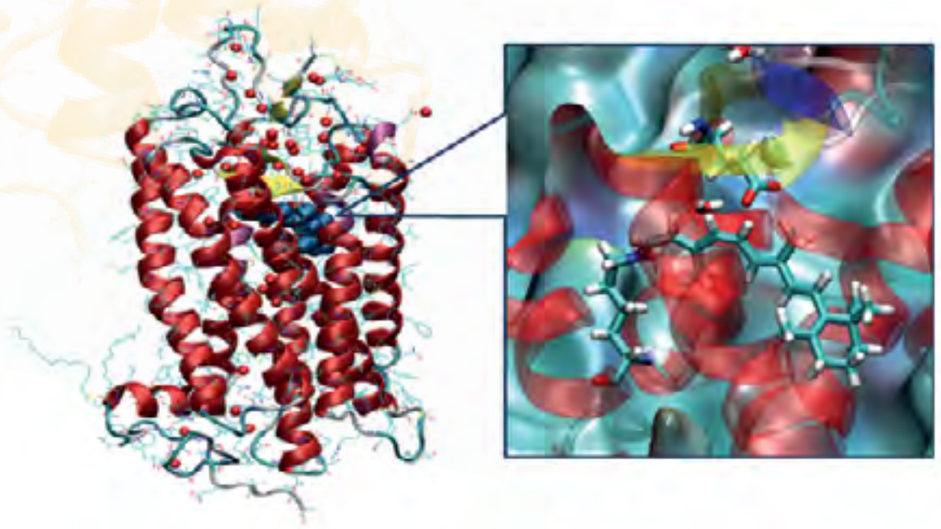


Figure 2: (Left) Dim-light visual pigment Rhodopsin, with the retinal chromophore inside (depicted in blue). (Right) The 11-cis PSB retinal chromophore covalently bound to the protein inside its binding pocket.

- We have identified higher electronically excited states of the isolated GFP chromophore anion in the previously unexplored UV region down to 210 nm [6]. By forming a dense manifold, these molecular resonances are found to serve as a doorway for very efficient electron detachment in the gas phase. Being resonant with the quasi-continuum of a solvated electron, this electronic band in the protein might play a major role in the GFP photophysics, where resonant GFP photoionization in solution triggers the protein photoconversion with UV light.

- Computationally, the largest run, ever performed using a highly correlated multi-reference electronic structure theory, has been executed in the XP-mode using 2,944 cores with a sustained performance close to 60% of the theoretical aggregate peak performance, ca 40 Teraflop/s. The calculations are aimed at understanding of a wavelength tuning mechanism in the retinal-containing visual proteins, where a single 11-cis PSB retinal chromophore is responsible for entire color vision (Fig. 2).

On-going Research / Outlook

Our results indicate the existence of efficient electron-nuclear coupling mechanisms in the photoresponse of biological chromophores. These non-adiabatic mechanisms are found to be dual, and, importantly, mode-specific. Numerous ways may be outlined by which the proteins use this electron-to-nuclei coupling to guide the photochemistry and the photophysics that lie behind their functioning. In the future projects, we plan to unravel the details of such non-adiabatic mechanisms, thus enhancing our understanding of the efficiency of light triggered processes in nature. Our findings will ultimately lay the groundwork for mode-selective photochemistry and photophysics in biological systems.

We also expect direct benefits from further tuning and optimizing the Firefly code, aiming at efficient large-scale excited-state and non-adiabatic calculations. This will expand a range of biologically-relevant systems, which will be feasible for high-level ab initio calculations, thus further pushing the limits of theory.

Acknowledgements

This work was supported by the RFBR (grant No. 14-03-00887). The results have been achieved using the PRACE-2IP project (FP7 RI-283493) resources LRZ's SuperMUC, HLRS's Laki, and RZG's MPG Hydra based in Germany.

References and Links

[1] **Granovsky, A.A.**
Firefly, http://classic.chem.msu.su/gran/firefly/index.html

[2] **Granovsky, A.A.,**
J. Chem. Phys. 134, 214113, 2011.

[3] **Granovsky, A.A.**
http://classic.chem.msu.su/gran/gamess/qdpt2.pdf

[4] **Granovsky, A.A.**
http://classic.chem.msu.su/gran/gamess/table_qdpt2.pdf

[5] **Bochenkova, A.V., Andersen, L.H.**
Faraday Discuss. 163, 297–319, 2013.

[6] **Bochenkova, A.V., Klaerke, B., Rahbek, D.B., Rajput, J., Toker, Y., Andersen, L.H.**
Phys. Rev. Lett., submitted, 2014.

contact: Anastasia Bochenkova, bochenkova@phys.au.dk

- Anastasia V. Bochenkova [1]
- Alexander A. Granovsky [2]

[1] Department of Physics and Astronomy, Aarhus University, Denmark

[2] Firefly project, Moscow, Russia

Applications

NIC Excellence Project 2013

Ab Initio Geochemistry of the Deep Earth

Most of what we know today about the formation and evolution of the Earth is derived from a combination of geophysical observations, geochemical information extracted from minerals or rocks, and laboratory experiments. An essential prerequisite to reconstruct the history of our planet is the knowledge of the thermodynamics and kinetics of geological processes in a wide range of relevant time and length scales. Chemical elements and their isotopes are redistributed most effectively at high temperatures and in the presence of silicate melts or aqueous fluids. Chemical transport in minerals and rocks is usually much slower. In some cases signatures of the time and the thermodynamic state of their formation can be stored over billions of years.

While melts and fluids play a very important role in geological processes, the investigation of their molecular structure, physical and thermodynamic properties remains challenging, especially under the extreme conditions of pressure and temperature prevalent in the Earth's interior. With the rising power of supercomputers, first-principles simulations based on quantum mechanics and density-functional theory [1] have become an increasingly powerful tool to complement experimental approaches in studying mineral-fluid or mineral-melt interactions. Such simulations are not only needed to interpret e.g. the complex fingerprint information obtained by various types of spectroscopy, but they also provide a unique link between chemical interaction, structure, and physical properties of the respective phases.

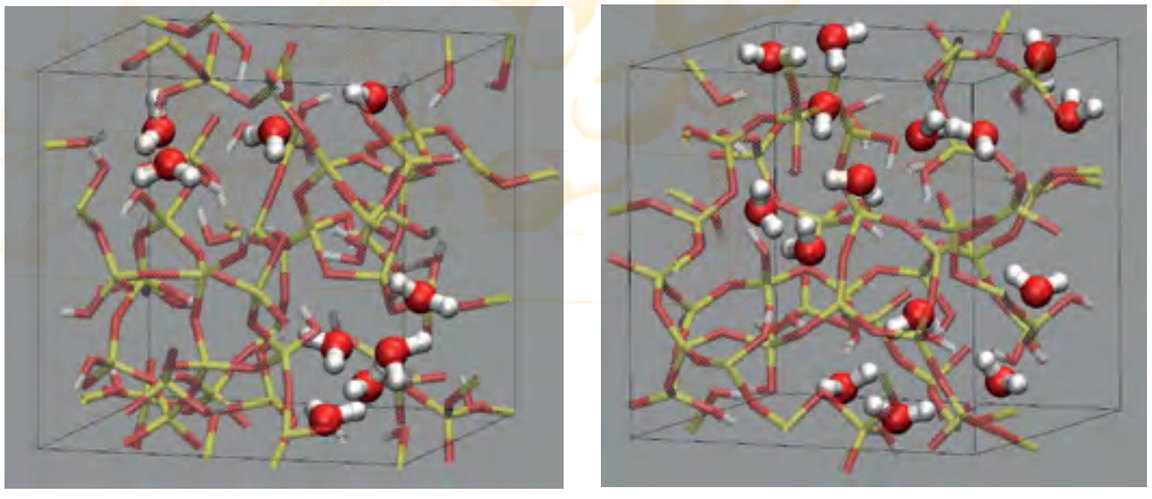


Figure 1: Structural evolution of an H_2O-SiO_2 fluid during isochoric cooling from 3,000 K (left) to 2,400 K (right) [2]. The increasing number of H_2O molecules (shown as balls and sticks) may be interpreted as a precursor of fluid-melt phase separation, which is also an important process in explosive volcanism.

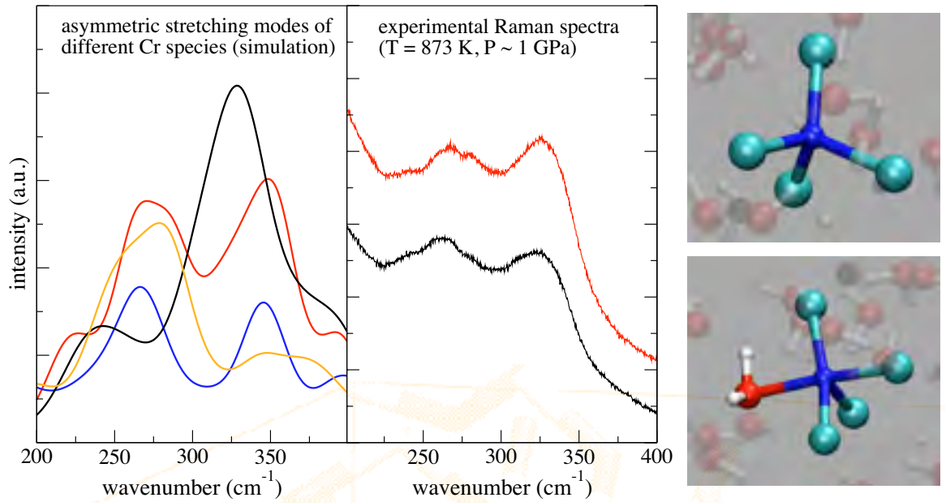


Figure 2: Vibrational mode spectra of chloridic Cr(III) complexes in aqueous solution from first-principles molecular dynamics simulations compared to experimental Raman spectra [5]. Also shown are two snapshots from the simulations.

Molecular Structure of Melts and Fluids at High Pressures and Temperatures

A first step towards a better understanding of the role of melts and fluids in geological processes from a molecular scale perspective is the development of realistic structure models for these disordered phases at relevant conditions. For this purpose, state-of-the-art experiments using methods such as x-ray diffraction, x-ray absorption, infrared absorption, or Raman spectroscopy need to be performed in situ to account for the continuous structural changes with variation of pressure, temperature, and/or chemical composition. However, there is no unique structure solution from the experimental data alone. Various molecular modeling approaches have been developed in the last decades to address this problem. But only with the feasibility of performing first-principles molecular dynamics simulations on supercomputers do we now have a method available to make realistic structure predictions for chemical complex systems at extreme conditions.

The change in the molecular structure of a fluid with temperature is illustrated in Fig. 1. Ab initio molecular dynamics simulations of an H_2O-SiO_2 fluid at temperatures of 3,000 K and 2,400 K at pressures of about 4 GPa show that the number of free H_2O molecules in the fluid increases with decreasing temperature and pressure [2], which eventually leads to a phase separation into a silica-bearing aqueous fluid and a hydrous silicate melt. The hydrolysis reaction, which transforms bridging Si-O-Si bonds into non-bridging Si-OH groups, can still be observed directly on the timescale of picoseconds at these high temperatures. At temperatures below about 2,000 K the rate of formation or breaking of strong Si-O

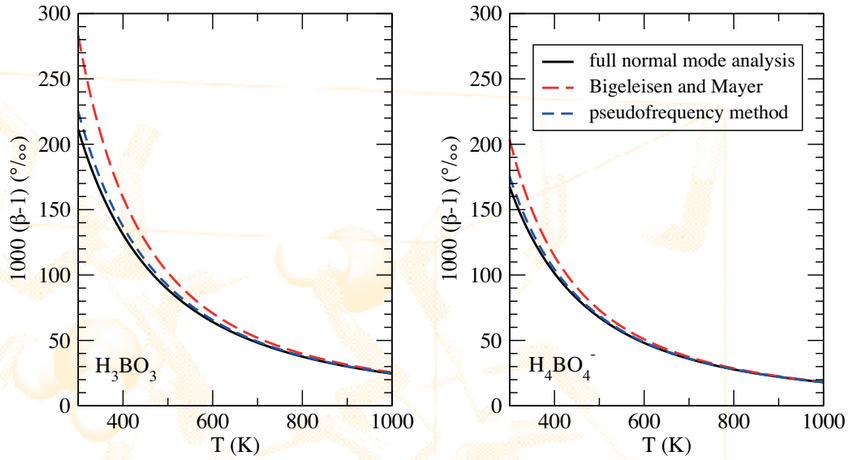


Figure 3: Comparison of $^{11}B/^{10}B$ isotope fractionation factors as a function of temperature for two molecules, H_3BO_3 and $H_4BO_4^-$ computed using the different methods described in the text (modified from [6]).

bonds is becoming too low, and acceleration methods such as constrained dynamics or metadynamics need to be employed to sample the equilibrium state. In systems with weaker bonds such as aqueous solutions containing mono- or divalent cations, the lowest temperature for direct observation of changes in the cation complexation decreases to lower than 1,000 K, which then covers relevant conditions of hydrothermal or subduction zone environments.

Depending on the objectives of a specific study we may be interested in the global structure of a melt or fluid as probed e.g. by diffraction methods, or in the atomic environment of a specific element that is present in the system with very low concentration. In the latter case, a site-sensitive spectroscopic method would provide a reference of the 'real system'. The most direct assessment of the predicted structure model is made by comparing experimental and computed diffraction patterns, vibrational or electronic excitation spectra. For this purpose we use the molecular dynamics trajectories in conjunction with various methods of theoretical spectroscopy [3]. Furthermore, our simulations support the interpretation of experimental spectra. For instance, application of a mode-projection technique [4] allows us to relate bands in experimental Raman spectra of high pressure / high temperature fluids to vibrations of individual cation complexes studied in the simulation (Fig.2) [5].

Equilibrium Stable Isotope Fractionation between Minerals and Fluids

Stable isotopes are widely used as geochemical tracers, which includes applications in deep Earth geochemistry as well as in paleoclimatology. The field of isotope geochemistry has been expanding greatly since modern analytical methods allow measuring variations in isotope concentrations in the sub-per mil range. From a thermodynamic point of view, isotope fractionation between different phases is caused by small differences in the Gibbs energy due to the mass-dependence of the normal mode vibrational frequencies. It is known that the heavy isotope preferentially fractionates into the stronger bonded environment. Using first-principles methods we are now able to make quantitative predictions of equilibrium stable isotope fractionation between different phases including minerals, melts, and fluids. While density-functional perturbation theory or other quantum chemical methods have been employed to compute fractionation factors for (simple) crystals and molecules by a number of groups, the explicit treatment of a fluid or melt phase remained almost unexplored. Major challenges are (1) the need of a reliable structure model for the fluid at relevant pressures and temperatures and (2) an efficient scheme to compute the fractionation factor from a large number of molecular dynamics configurations representing a complex disordered material.

Within our NIC project, we have developed an approximate method that predicts mineral-fluid isotope fractionation with accuracy similar to experiment [6]. As described above, the structure of the fluid is sampled by first-principles molecular dynamics simulation. The fractionation factor, which is expressed as a reduced partition function ratio, is then derived considering the local nature of the fractionation process. A conventional full normal mode analysis would require a structural relaxation of all atoms in the simulation cell followed by computing harmonic frequencies. This is not only computationally extremely expensive but it also changes the structure of the fluid. It turns out that at high temperatures it is sufficient to relax only the isotopic atom and its nearest neighbors, and then to compute the force constants acting on the isotopic atom in the three Cartesian directions (Bigeleisen and Mayer approximation). If these force constants are transformed into pseudo-frequencies (generally these are not normal modes) and the latter are inserted into the reduced partition functions, the obtained fractionation factors are very similar to those obtained from a full normal mode analysis (Fig. 3). After having demonstrated the accuracy of our new method for Li and B isotopes we are now confident to make reasonable predictions for other important light isotopes such as Si and to shed new light on fractionation processes, e.g. during Earth's core formation.

Towards Predictive Modeling of Element Partitioning and Mineral Solubility

In a next step, we have started to address the problem of element re-distribution during fluid-rock or melt-rock interaction. This is an important field in geochemistry ranging from the use of trace elements as geochemical tracers to element mobilization in hydrothermal

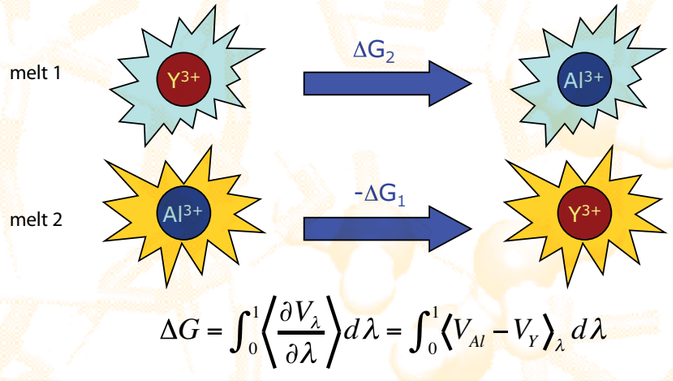


Figure 4: Computing the change in Gibbs energy ΔG using the alchemical change method. After adding the two partial transmutation reactions, the trace element (here Y^{3+}) partition coefficient between two phases can be extracted assuming a partition coefficient close to unity for the major element (Al^{3+}).

fluids during ore-forming processes. First-principles prediction of element partitioning between different phases is conceptually more complex than isotope fractionation. For computing Gibbs energy differences the changes in the chemical interactions upon element substitution need to be accounted for. Furthermore, Gibbs energies cannot be derived directly from a single molecular dynamics trajectory. One promising approach is to use the alchemical change method of thermodynamic integration as illustrated in Fig. 4, which we recently applied in a study of the melt-dependence of the trace element partitioning of Y^{3+} between silicate minerals and melts [7]. In a number of molecular dynamics simulations a single particle in the simulation cell is transmuted from one element to another (here from Y^{3+} to Al^{3+}) by changing the interaction potential V through variation of the integration parameter λ in steps from 0 to 1. After performing this transmutation in the two phases of interest the exchange coefficient of the two elements between the two phases can be derived. While in our pilot study [7] we used classical interaction potentials we are currently exploring the possibility to predict realistic element partition coefficients from first principles simulations. This challenging goal is computationally very demanding and can only be achieved by using supercomputer facilities.

Acknowledgements

This work has been funded by the Deutsche Forschungsgemeinschaft within the Emmy-Noether program (grant no. JA1469/4-1). Simulations were performed on the supercomputers JUQUEEN and JUROPA at JSC thanks to allocated computing time within NIC project HP015.

References

[1] **Marx, D., Hutter, J.**
Ab-initio molecular dynamics: Basic theory and advanced methods. Cambridge University Press, 2009.

[2] **Spiekermann, G.**
PhD thesis, FU Berlin, 2012.

[3] **Jahn, S., Kowalski, P.M.**
Rev. Mineral. Geochem. 78, 691-743, 2014.

[4] **Spiekermann, G., Steele-MacInnis, M., Schmidt, C., Jahn, S.**
J. Chem. Phys. 136, 154501, 2012.

[5] **Watenphul, A., Schmidt, C., Jahn, S.**
Geochim. Cosmochim. Acta 126, 212-227, 2014.

[6] **Kowalski, P.M., Wunder, B., Jahn, S.**
Geochim. Cosmochim. Acta 101, 285-301, 2013.

[7] **Haigis, V., Salanne, M., Simon, S., Wilke, M., Jahn, S.**
Chem. Geol. 346, 14-21, 2013.

- Sandro Jahn
- Daniela Künzel
- Johannes Wagner

Deutsches GeoForschungs-Zentrum GFZ, Potsdam

contact: Sandro Jahn, jahn@gfz-potsdam.de

Nonlinear Response of Single Particles in Glassforming Fluids to External Forces

In "active micro-rheology" the response of a tagged particle in a fluid to an external force f that acts only on this particle is measured [1]. This new experimental method can be applied to colloidal and biological systems, where it can be realized, e.g., by using a combination of magnetic tweezers and confocal microscopy. Particularly interesting is the use of active micro-rheology to study glass-forming fluids, since these systems show a strong non-linear response even for very small external forces [2], and understanding the glassy freezing of such systems still is one of the grand challenge problems of condensed-matter physics [3].

In order to provide an understanding how such experiments have to be properly analyzed and interpreted, a faithful model of this approach has been implemented via large scale computer simulations on the JUROPA supercomputer of the NIC/JSC. As a model glass-former, a binary mixture of colloidal particles that interact with screened Coulomb potentials (so-called "Yukawa potentials") has been used. In the Non-equilibrium Molecular Dynamics (NEMD) simulation, a single particle is pulled with a constant external force of strength f that acts in x-direction. The simulation box contains altogether 1,600 particles; periodic boundary conditions are applied in all directions to make the system representative for a macroscopic colloidal dispersion. The dynamics of this interacting many-particle system is simulated by numerically solving Newton's equations of motion, as usual; however, it is essential to use a so-called "dissipative particle dynamics" (DPD) thermostat to keep the chosen temperature constant, since the extra heat put into the system via the additional friction that the pulled particle creates in the system needs to be removed. The DPD thermostat is constructed such that it is compatible with the proper hydrodynamic equations and their underlying conservation laws.

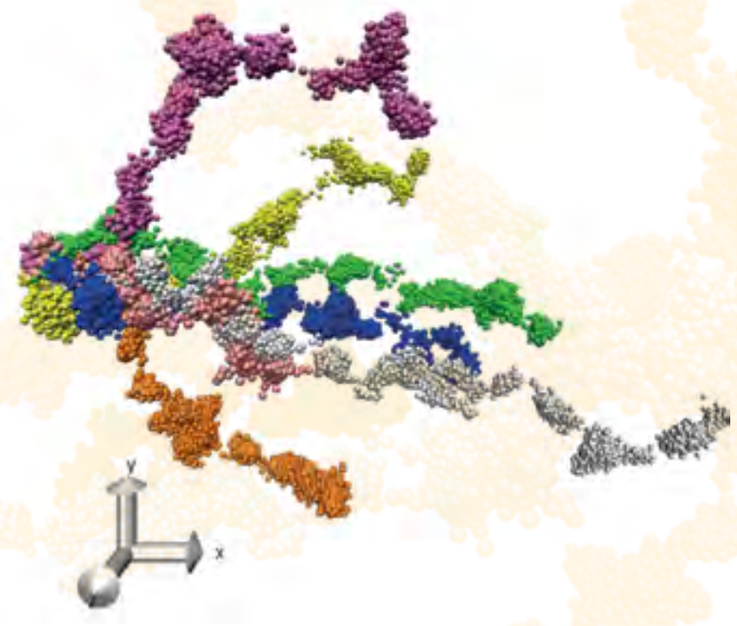


Figure 1: Snapshot of typical trajectories of a pulled A-particle in the binary AB-mixture (containing 800 particles of type A and 800 of type B) for a temperature T = 0.14 slightly above the glass transition temperature and a force f = 1.5 (the force is measured in units of $k_B T/d_{AA}$, where k_B is Boltzmann's constant and d_{AA} the diameter of A particles; temperature is measured in units of the energy scale ε_{AA} of the Yukawa potential between A-particles, $V_{\alpha\beta}(r) = \varepsilon_{\alpha\beta}(d_{\alpha\beta}/r) \exp[-6(r/d_{\alpha\beta}-1)]$, with α,β = A,B). The trajectories are drawn such that they all start at the same initial point.

When one follows the trajectories of the pulled particles in order to study the time evolution of their mean square displacements (note that the latter are anisotropic, directions parallel and perpendicular to the direction of the pulling force are not equivalent), one encounters huge sample-to-sample fluctuations from run to run, and therefore huge computational resources need to be invested to obtain physically meaningful results (by sampling thousands of trajectories over a significant range of time). However, these huge fluctuations are one of the central subjects of the study, since they reflect the so-called "dynamical heterogeneity" [3] of the supercooled fluid (Fig. 1). Dynamical heterogeneity means, that the fluid has a structure like an irregular mosaic, where in different "stones" of the mosaic the local mobility of the particles may differ from the mobility in a neighboring regime by orders of magnitude! Such differences in mobility of glass-forming fluids can be caused by tiny density fluctuations already. The trajectories of the pulled particles then exhibit a pearl-necklace type appearance (Fig. 1): When the particle is "caught" in a cage formed by its neighbors, a cooperative rearrangement of a rather large surrounding region may be necessary until the particle can "escape from the cage". These motions in the cage are reflected in a pearl-like structure of the trajectory. Conversely, when the particle passes a region of higher mobility, the force can pull it through such a region rather fast, causing an almost linear string-like piece of the trajectory.

When one follows each trajectory long enough, one clearly can identify a velocity v of the steady state that is established, and identify a friction coefficient ξ of the pulled particle via $\xi = f/v$. This friction coefficient is found to be a strongly nonlinear function of the force f at low temperatures (Fig. 2a). Particularly illuminating is this behavior when one describes it in terms of a so-called "Peclet number" Pe*, see Fig. 2b. One finds that at low enough temperatures ($T < 0.18$) and intermediate forces outside the linear response regime the data can be scaled such that they all fall on a master curve (Fig. 2b, inset). This implies that there holds a "force-temperature superposition principle" [4], similar to the well-known "time-temperature superposition principle" [3] for relaxation functions of glass-forming systems in thermal equilibrium.

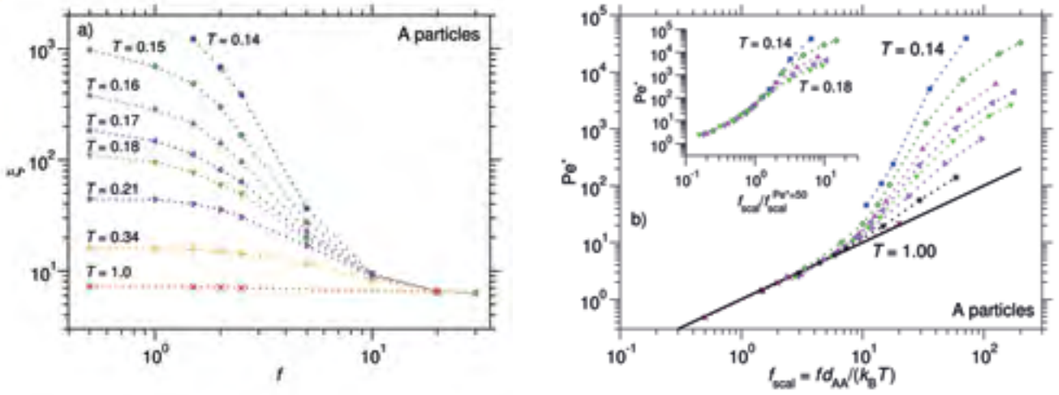


Figure 2: a) Friction coefficient ξ for A particles and different temperatures T as function of the force f. b) Peclet number $Pe^* = f D_{eq} d_{AA} / \xi$ plotted vs. the scaled force $f_{scal} = f d_{AA}/k_B T$. The solid line indicates the linear response behavior, for which $Pe^* = f_{scal}$ holds (D_{eq} is the self-diffusion coefficient of the A particles in equilibrium). The inset shows Pe^* as a function of the ratio $f_{scal} / f_{scal}^{Pe^*=50}$. From Ref. [4].

A very interesting behavior is also detected when one studies the time-dependent mean square displacements (MSDs) of the pulled particles (Fig. 3). For $f = 0$, there are three regimes: at small times, there is a ballistic regime ($MSD \propto t^2$), when the particle's motion is not yet constrained by its neighbors. The "cage effect" (i.e., a particle is "caught" by a cage formed by its neighbors) shows up by almost horizontal plateaus in the MSD vs. t curves. When the particle is released from the cage (which happens by cooperative rearrangement of the particles in its neighborhood), ordinary diffusive motion ($MSD \propto D_{eq}\, t$) sets in. When a force f is applied, we see that the "lifetime" of the cage is reduced (the more the stronger the force), and afterwards a super-diffusive motion in the direction along the force sets in (while in the perpendicular directions still ordinary diffusive behavior occurs). The anomalous behavior of the MSDs is also reflected in the distribution functions of the displacements, the so-called van Hove correlation functions (see [5,6] for details). Also theoretical concepts (such as the extension of the so-called "mode coupling theory" to nonlinear response of glass-forming fluids) can be stringently tested [6]. Thus the studies carried out in this research project yield important ingredients towards the theoretical understanding of dynamical heterogeneity in glass-forming fluids.

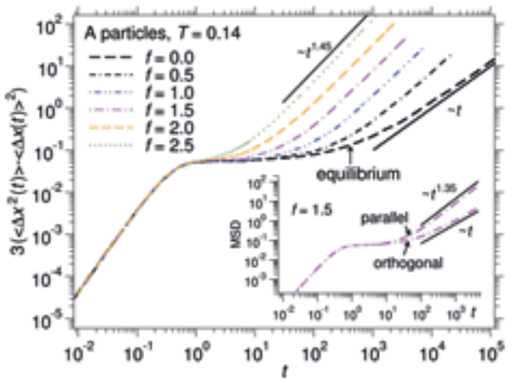


Figure 3: Mean-squared displacement (MSD) of A particles in the direction of the force for different values of its strength f, as indicated, for T = 0.14. The straight lines show exponents for ordinary diffusion ($\propto t$, observed for $f = 0$) and for super-diffusive behavior ($\propto t^{1.4}$). The inset shows a comparison between MSDs for the parallel and perpendicular direction for $f = 1.5$. From Ref. [4].

Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft (DFG), Projects No. SFB TR6/A5 and No. FOR 1394/P8, and by a generous grant of computer time at the JUROPA supercomputer of the Jülich Supercomputer Centre (JSC) provided via the John von Neumann Institute for Computing (NIC).

References

[1] **Squires, T. M., Mason, T. G.**
Annu. Rev. Fluid Mech. 42, 413, 2010.

[2] **Habdas, P., Schaar, D., Levitt, A. C., Weeks, E. R.**
Europhys. Lett. 67, 477, 2004.

[3] **Binder, K., Kob, W.**
Glassy Materials and Disordered Solids: An Introduction to Their Statistical Mechanics Rev. Ed. (World Scientific, Singapore, 2011).

[4] **Winter, D., Horbach, J., Virnau, P., Binder, K.**
Phys. Rev. Lett. 108, 028303, 2012.

[5] **Winter, D., Horbach, J.**
J. Chem. Phys. 138, 12A512, 2013.

[6] **Harrer, C., Winter, D., Horbach, J., Fuchs, M., Voigtmann, T.**
J. Phys.: Condens. Matter 24, 464105, 2012.

contact: Jürgen Horbach,
horbach@thphy.uni-duesseldorf.de

- Kurt Binder[1]
- Jürgen Horbach[2]
- Peter Virnau[1]
- David Winter[1]

[1] Institut für Physik, Johannes Gutenberg-Universität Mainz

[2] Institut für Theoretische Physik II, Heinrich Heine Universität Düsseldorf

Applications

Is there Room for a Bound Dibaryon State in Nature? – An ab Initio Calculation of Nuclear Matter

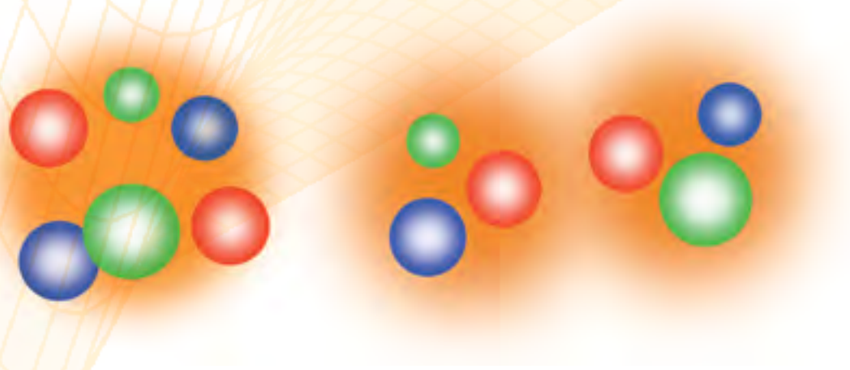


Figure 1: A schematic of a dibaryon (left) and two loosely bound baryons (right). All nuclear matter is made up of quarks and gluons. They are governed by the laws of QCD.

What are the fundamental properties of nuclear matter? How does it behave under extreme conditions? – These are some questions that ab initio calculations of hadrons and nuclear matter using numerical lattice techniques are trying to answer. Focussing on one, we ask: Is there room for a bound dibaryon state in nature?

The dibaryon (a 6-quark state) was predicted over 30 years ago in a model calculation [1]. So far, however, its detection has eluded experimental physicists, and theoretical studies remain inconclusive. It is therefore interesting to take the step from the well determined model-independent numerical calculations of lattice Quantum Chromodynamics (QCD) for single baryons (3-quark states, such as the proton) to dibaryons – thus paving the way to ever more complicated nuclei. We study the spectrum of (di)baryons in terms of correlation functions that determine the properties of a particle state, such as its mass and whether it is a loosely bound system of individual particles or a tightly bound state (Fig. 1). The interpolating operators involved in correlation functions are related to the particle's wavefunction, which is mostly determined by the individual (valence) quarks it consists of. This is the realm of QCD and our goal is the numerical ab initio lattice QCD calculation of dibaryon properties.

Quantum Chromodynamics (QCD)

At the core of all nuclear matter lies QCD. It is the field-theoretical description of the fundamental strong force, which determines the interactions between quarks and gluons, i.e. the constituents of hadrons.

QCD has two distinctive properties: asymptotic freedom at high energies, and confinement at low energies. Asymptotic freedom corresponds to a small coupling constant in the theory, and in this regime perturbation theory gives a highly precise description of the physical phenomena. Nuclear matter, however, is a low-energy phenomenon and a different approach must be used to explore hadronization, i.e. the formation of bound states and matter from individual quarks. Lattice QCD is a first-principles approach that has exactly this capability.

Lattice QCD

In lattice QCD simulations, the properties of baryons and nuclei can be calculated by studying correlation functions of local operators. These expectation values are linked to functional integrals of QCD that are evaluated stochastically using Monte Carlo methods. In addition, every 'measurement' must be performed on large ensembles of field configurations generated by a Markov process and distributed according to the QCD action. The extremely large number of variables in these simulations makes it impossible to use conventional computing resources. At the same time, lattice QCD simulations are well suited for massively parallel computers, as they involve local interactions and require the inversion of large sparse matrices. Lattice QCD is thus well suited for high performance computing (HPC).

To set up a lattice QCD calculation, spacetime has to be discretized on a hypercubic lattice. This induces systematic effects that need to be considered. These are the finiteness of the introduced lattice spacing between the spacetime points at which measurements are made, and the finite extent of the lattice. Effects due to the breaking of symmetry (e.g. rotational) down to a subgroup introduced by the hypercubic lattice must also be considered. Moreover there is the added complication that the simulations become increasingly expensive as the physical limit is approached, specifically when tuning the light quarks to their physical values. To

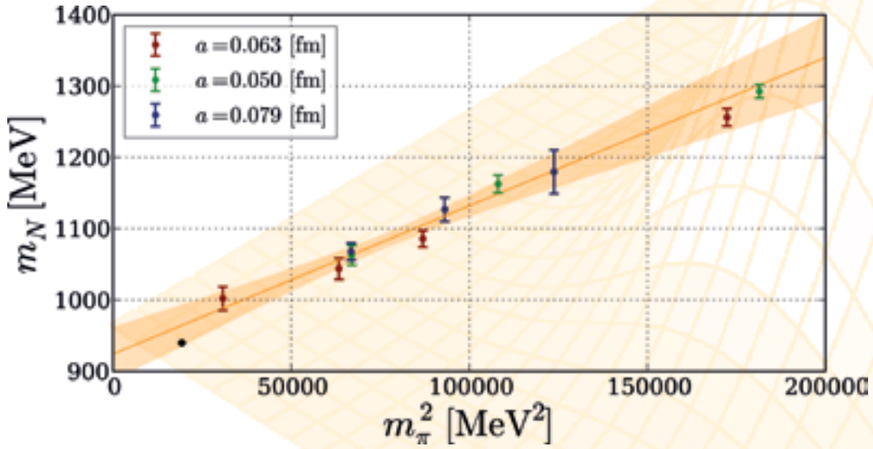


Figure 2: In lattice QCD, several calculations are performed at different unphysical quark masses (colored points). In this way, the systematic effects can be taken into account and results can be safely extrapolated to the point of the experimentally measured quark masses. The black point shows the experimentally determined point.

handle these effects, multiple series of measurements are usually made at several unphysical quark masses, several different lattice spacings and lattice volumes. Through closely monitoring the lattice QCD results, a safe and reliable extrapolation to the physical limit may be performed. In the case of the single proton, this has been performed extensively so that a reliable extrapolation to the physical point with controlled systematics can be made (Fig. 2). However, in the much more complicated case of nuclear matter, i.e. nuclei of mass number $A \geq 2$, this type of calculation is still in its infancy. As a consequence, the following initial calculation is on a lattice volume of 64×32^3 (time × space^3) points, for a pion mass of 450 MeV and a spatial extent of 2 fm. A larger and more physical ensemble of ($m_{\pi,}$=320 MeV, 96×48^3, and 3 fm) will shortly be analyzed on JUQUEEN.

Dibaryons

Hadrons are composite particles, consisting of quarks bound by the strong force. There are 6 quarks in total. However we simulate only the 3 lightest quarks: the up (u), down (d) and strange (s) quarks (in ascending order of mass). There are two families of hadrons that have been observed in nature. The first family is that of quark-anti-quark states called mesons, with the pion being its most prominent member. The second is that of 3-quark states referred to as baryons, of which the stable proton (quark content uud), the long lived neutron (udd, with a lifetime of ~15 minutes) and Λ-baryon (uds) are members. In addition, the dibaryon is a proposed 6-quark state (Fig 1). The H-dibaryon is a state with strangeness quantum number –2 and quark content udsuds. This state was predicted in the 1970s, and an indication that such a state exists would be that it is energetically favored over two distinct baryons, i.e. if the H-dibaryon mass is less than twice the single Λ-baryon mass.

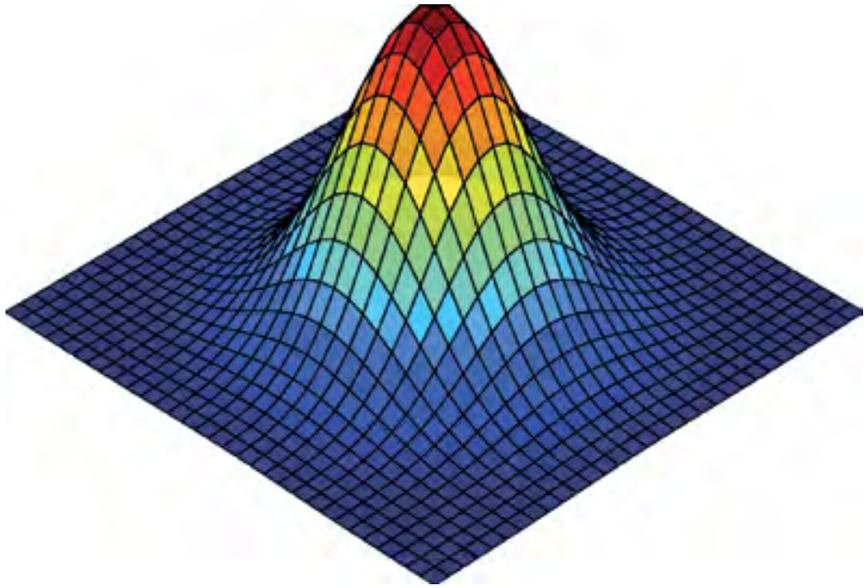


Figure 3: An optimized spatial profile of quark sources used in our calculation of the H-dibaryon state [2].

Numerics

Our preliminary findings indicate that, in order to resolve the mass difference between two individual Λ-baryons and a true dibaryon state, a statistical precision of <1% for the masses of the individual and dibaryon states is required. This means that an extremely large number of individual lattice QCD measurements are required, of order 30,000 (corresponding to approximately 3.8 Mcore-hrs on JUQUEEN for our chosen parameters).

To set up a calculation of the H-dibaryon, a QCD interpolating operator that describes it must be used. This is achieved by forming a combination of quark fields so that the resulting interpolating operator has the correct quantum numbers to be related to the H-dibaryon wavefunction. These quark fields then have to be connected ('contracted') to form a measurable correlation function. Numerically, any single lattice QCD measurement of a hadron mass thereby consists of three main parts: the calculation of quark propagators (matrix inversions); techniques used to obtain a good overlap with the state of interest, which requires some tuning of the operator (Fig. 3); and the contraction of the quark propagators to form a hadronic correlation function. For a baryon, the first two are the most expensive. However, due to the larger number of combinations of quark propagator contractions that enter the dibaryon correlation function, stemming from the additional 3 quarks, all three steps are equally significant for a dibaryon measurement, thus making it more computationally demanding.

The number of combinations of quark propagators required for a (di)baryon calculation is equal to the factorials of their respective quark content $N = N_u! N_d! N_s!$, which is 2 for the proton (uud), 8 for the H-dibaryon (udsuds), and 36 for a deuteron (uududd). For other nuclei of mass number $A \geq 2$ this rapidly increases. Consequently, the H-dibaryon system is much more complex and demanding than the single proton. To improve the efficiency of the calculation, we employ two techniques: a 'blocking' algorithm and a 'truncated' solver.

Blocking Algorithm

To tackle the large number of contractions, it is convenient to use the so-called 'blocking' algorithm [3] to handle and simplify the potentially large number of combinations of quark propagators (Fig. 4). This involves a pre-contraction of the quark propagators at the source of the correlation function, so that the pre-contracted 'block' can be re-used to efficiently combine the quark propagators at the sink. This removes the need for a new source contraction for each combination.

Truncated Solver

Quark propagators are calculated through the inversion of a large sparse matrix using an iterative solving procedure. The truncated solver method [4] utilizes a less precise stopping condition (tolerance of 10^{-3} compared to the usual 10^{-10}) for the inversion. There is of course an associated bias with respect to a high precision solve. However this can be corrected for, which allows many more measurements to be made in a given time compared to the usual high-precision method and results in at least a 20% gain in statistical precision.

The truncated solver and the aforementioned blocking algorithm are two very new and modern technical developments that are key to making this calculation possible. With these improvements, the calculation may be achieved using existing HPC resources in a sensible time-frame and has become feasible for lattice ensembles large enough to model the dibaryon without significant distortion due to the finiteness of the lattice spacing and with lattice parameters approaching those of the physical situation.

Results

We show results gathered on a total of ~16,000 measurements [5] obtained in 1.9 Mcore-hrs of computing time on JUQUEEN. The mass of the H-dibaryon was determined using the exponential decay of the correlation function. For large time separations the ground state dominates the correlation function and excited states are exponentially suppressed. The ground state mass is most easily computed from taking the logarithmic derivative of the correlation function to determine an 'effective' mass. The ground state mass can then be determined by performing a fit in the plateau region of the effective mass (seen in Fig. 5) where it is independent of the time separation. However, this method is only suited for a determination of the ground state. A more sophisticated method is required for a study of the excited states. For instance, we employ the method of solving a generalized eigenvalue problem (GEVP). The choice of interpolating operator used to measure the H-dibaryon is not unique, and one is free to choose any wavefunction that gives the correct quantum numbers. Depending on how well a given operator describes the actual physical state of the H-dibaryon, the correlation function will have a better or worse overlap with the true H-dibaryon wavefunction. Solving the eigenvalue problem requires a set of such differently overlapping operators and provides the mass eigenvalues of the operator matrix, which allows the corresponding state's mass to be retrieved. Our analysis uses a set of 4 trial wavefunctions to extract the ground and first excited states of the H-dibaryon mass spectrum. The preliminary results in Fig. 5 show that with the present statistical accuracy of ~4%, the H-dibaryon mass is consistent with that of two individual particles. However for a truly conclusive result the statistical error on the dibaryon data must be significantly reduced.

Conclusions

The study of nuclear matter from ab initio lattice QCD calculations is still in its infancy and requires top-tier HPC resources for its study. Our first analysis shows a tendency of the H-dibaryon to be unbound. However, the statistical error needs to be substantially reduced. Additionally, the calculation has to be repeated on a series of lattice ensembles with ever more physical, and hence expensive, parameters. Even at this early stage the results do however give insights into the nature of light nuclei, specifically the H-dibaryon [6], that are beyond the reach of other non-ab initio approaches. With the continued investment of HPC resources on platforms such as JUQUEEN we will be able to answer the question: is there room for a bound dibaryon state in nature?

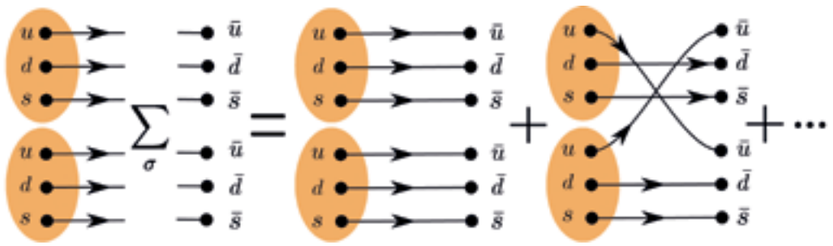


Figure 4: The dibaryon is a much more complex system than the single baryon and efficient algorithms have to be used to handle the large number of quark contractions [5]. There are 8 terms in total for the udsuds dibaryon.

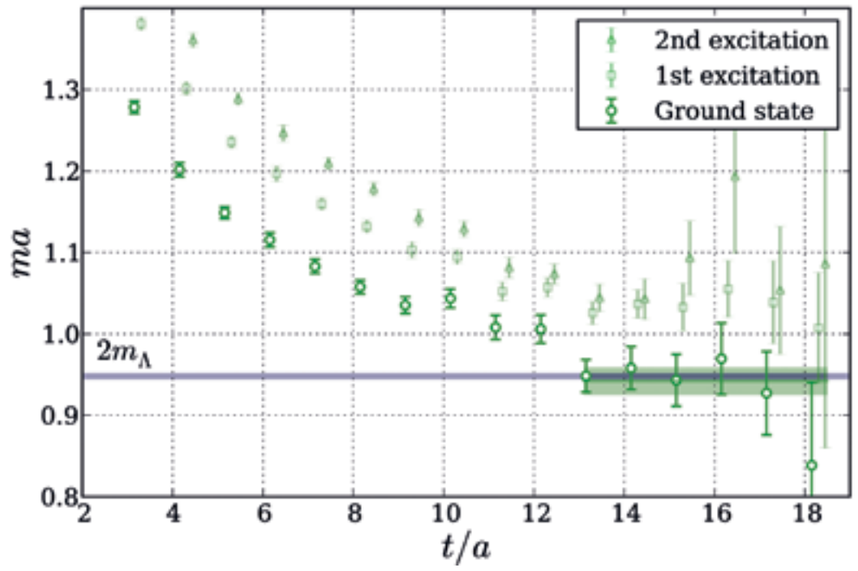


Figure 5: The GEVP of three interpolating operators gives the first three mass states of the H-dibaryon particle spectrum. If the lowest mass (ground state) lies below the threshold of two individual Λ particles (blue band), it is the energetically favored state [5].

Acknowledgements

The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time for a GCS Large Scale Project on the GCS share of the supercomputer JUQUEEN at Jülich Supercomputing Centre (JSC). This work was supported by the DFG via SFB 1044 and grant HA 4470/3-1.

References

[1] **Jaffe, R.L.**
Phys. Rev. Lett. 38 (1977) 195.

[2] **von Hippel, G.M., Jäger, B., Rae, T.D., Wittig, H.**
JHEP 1309 (2013) 014.

[3] **Doi, T., Endres, M.G.**
Comput. Phys. Commun. 184 (2013) 117.

[4] **Blum, T., Izubuchi, T., Shintani, E.**
Phys. Rev. D88 (2013) 094503.

[5] **Francis, A., Miao, C., Rae, T.D., Wittig, H.**
PoS (LATTICE 2013) 440, arXiv:1311.3933, 2013.

[6] **Walker-Loud, A.**
PoS (LATTICE 2013) 013, arXiv:1401.8259, 2013.

contact: Hartmut Wittig,
wittig@kph.uni-mainz.de

- Hartmut Wittig
- Thomas D. Rae
- Anthony Francis

Johannes Gutenberg
Universität Mainz

CoolEmAll–Energy Efficient and Sustainable Data Centre Design and Operation

Although IT hardware developers increase the energy efficiency of new systems, the global overall energy consumption of IT infrastructures rises year by year driven by the dramatically increasing usage of IT services covering all areas in our dailies life. The need of new strategies to reduce or at least minimize energy consumption is obvious. The European Commission funded project CoolEmAll has taken a holistic approach to the complex problem of how to make data centres more energy and resource efficient IT infrastructures by developing a range of tools to enable data centre designers, operators, suppliers and researchers to plan and operate facilities more efficiently.

The main concept for the project as shown in Fig. 1 includes the different aspects of data centre provisioning. Three different fields of input data are identified, first as major issue, the composition of hardware from compute nodes up to the room and infrastructure like cooling and power provisioning for which we used a modular approach (Data centre Efficiency Building Blocks, DEBBs) allowing to specifying parts of the whole systems independently and then combing them. Secondly the specification of the cooling mechanisms and their efficiencies is part of the base for the Simulation, Visualization and support toolkit (SVD Toolkit). The last piece is the description of the software running on the systems containing detailed information about application characteristics and the workload distribution. Based on this the SVD Toolkit can start the simulation process beginning with load distribution, calculating energy consumptions for all systems as input for the airflow and heat simulation and visualize the results and problems with the selected setup can be found like airflow recirculation or increased temperatures beyond limits. By repeating the simulation process with different setups an optimized configuration can be found where the energy consumption is minimized.

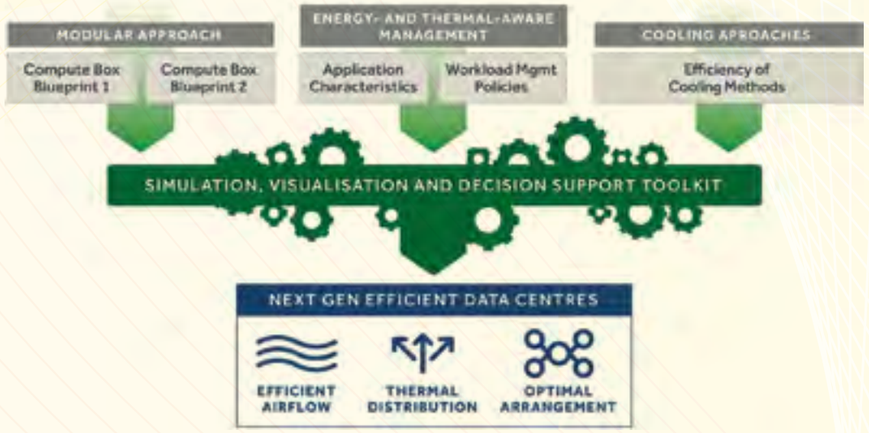


Figure 1: CoolEmAll main concept

The Data centre Efficiency Building Blocks (DEBBs) are introduced to allow the specification of separate blocks or sections of a data centre independently (i.e. node, enclosures, cooling devices, storage, …), since specifying the whole data centre in a monolithic way leads to very complex structures and increased effort to maintain the specification. DEBBs can also be created by the manufacturer of each product itself and then combined by the data centre designer or provider. DEBB are used to decouple independent section of a system.

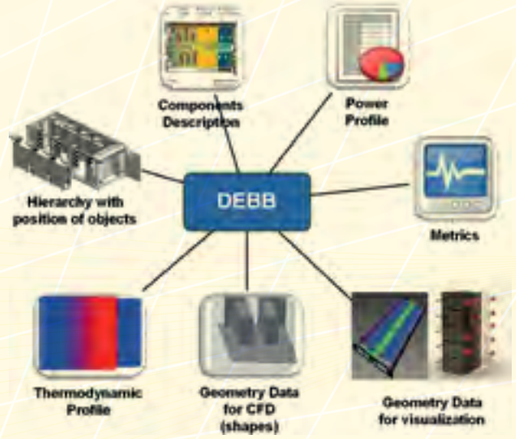


Figure 2: DEBB composition

Each DEBB represents a section of the system with either a certain functional behavior (like cooling, power provisioning, centralized infrastructure, etc.) or a more or less homogeneous part of the system (i.e. one cluster, one often used Rack, etc. with all necessary descriptions and details . As shown in Fig. 2 this includes Geometry information (3D model) describing in detail the physical structure used as input for the CFD simulations and for visualization, power

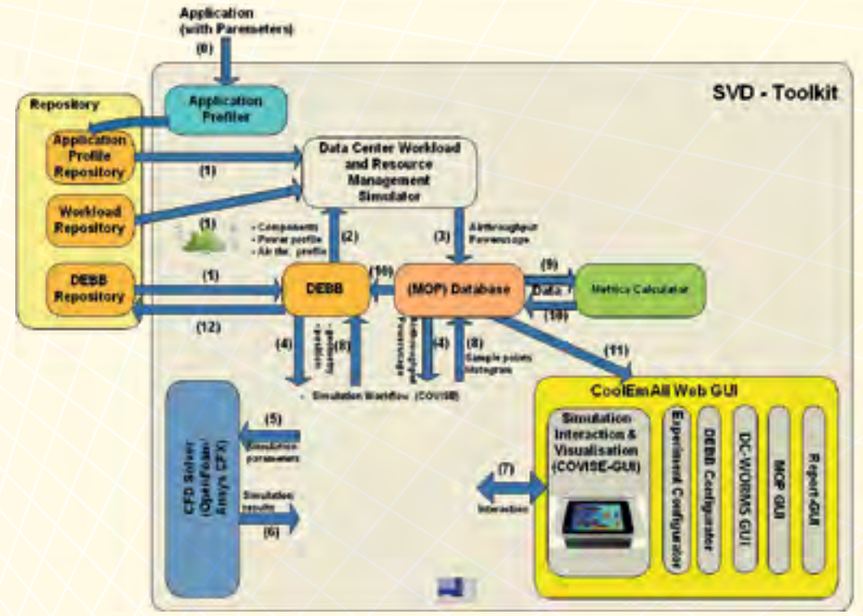


Figure 3: SVD Toolkit architecture

profiles describing the power consumption for any part depending on different operational states and loads, metrics describing which metrics are available for the specified device, component descriptions similar to fact sheet with common parameters for all devices and the hierarchy which represents a specific instance of the system, a whole data centre or only a node depending on the target to be achieved and can be seen as the glue between all the other pieces for the overall system. The hierarchy is the entry point and describes the composition of all devices

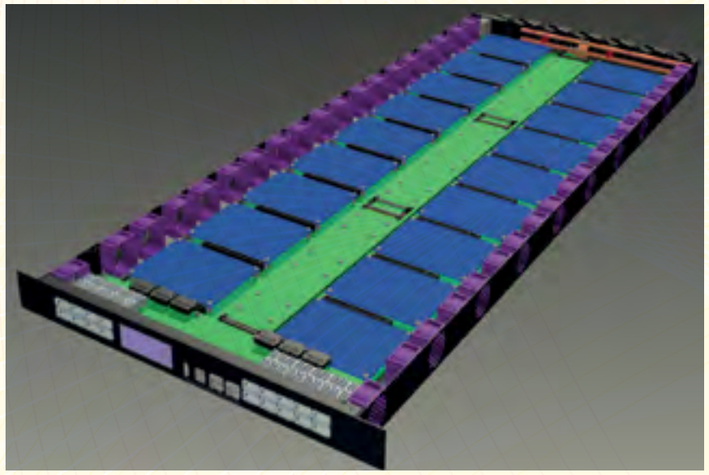


Figure 4: RECS Box 3D design

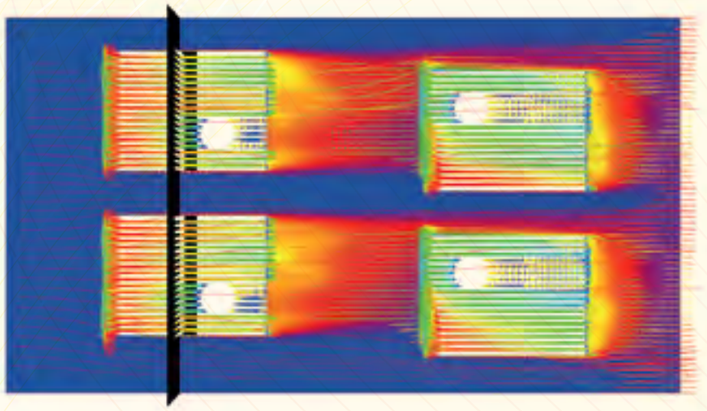


Figure 5: Simplified RECS heat distribution

and sub sections and refers to their corresponding component description, geometrical data or profiles. This modular approach also allows to specify different section on different levels of detail, depending on the need for the following simulation, so a well known system which is in the focus of the simulation can be specified more detailed than surrounding systems which are not homogeneous and equipped different so effort for specifying them also in detail would be to high in comparison to the relevance. It is also easy to replace parts of the DEBB description which might happen in case of an updated version of the description, when single devices are replaced or a system goes out of production. For the design phase this easy replacement is also useful to be able to compare different compositions without creating them new from scratch.

The SVD Toolkit (Fig. 3) developed in the project handles all simulations needed from application and workload simulations starting with the Data Center Workload and Resource Management System (DCWoRMS), the CFD simulations with either OpenFoam or Ansys CFX within the Collaborative simulation and Visualization Environment (COVISE). The simulation workflow definition together with the parameter and configuration setup as well as the visualization of the results is included in the CoolEmAll web GUI.

DCWoRMS uses the application profiles covering the applications behavior on different loads and the predicted or planned workload distribution to simulate mainly the usage rates of all components and as result the energy consumptions. This information can be used to sum up the overall capacity

utilization and optimization directly by reducing the provided hardware. For CoolEmAll the reduction of the overall energy consumption if far more important, so the energy values are used for the second simulation step, the airflow simulations. Within the project two types of simulations were used covering different issues. The first was the simulation of single enclosures as nodes or node groups like blade centers etc. This allows hardware manufacturer to optimize the layout of these devices by relocate openings or fans, changing the surface of heat sinks to improve the airflow and the dissipated heat resulting in reduced temperatures or lower power consumption for fans. Additionally the cooling in enclosures can be optimized by changing the order of the nodes in the airflow like for the RECS Box (Resource Efficient Computing and Storage in Fig. 4 and a simplified version for the simulation with only four nodes in Fig. 5) or changing in homogeneous environments the node to workflow assignment. For data centre designer and providers the other scenario is more important, the simulation of the whole server room including the cooling infrastructure. Since this is at a quite larger scale a simulation with the same granularity as before will last much longer producing results with lower accuracy limited by the fact the real world cannot be described identically, for example cables may influence airflows and therefore the accuracy will be not so high compared to the time and energy consumed for the simulation. Therefore all values are only used for a whole rack and the structure of the racks does also not cover any details within–it is taken as black box. This allows a fast and precise enough simulation on the room so that the designer and providers then can optimize the room layout, hardware selection and cooling setup to reduce the overall consumed energy.

With the SVD Toolkit developed within CoolEmAll data centre's energy consumption can be optimized by combining different strategies. Selecting other applications or hardware is often not appropriate but changes in the workload distribution should not lead to different results or throughput. Nevertheless optimizing the rack placement for rooms or changing the temperature slightly may decrease the cooling effort considerable. As shown in Fig. 6 (worse/good case) the changes in the

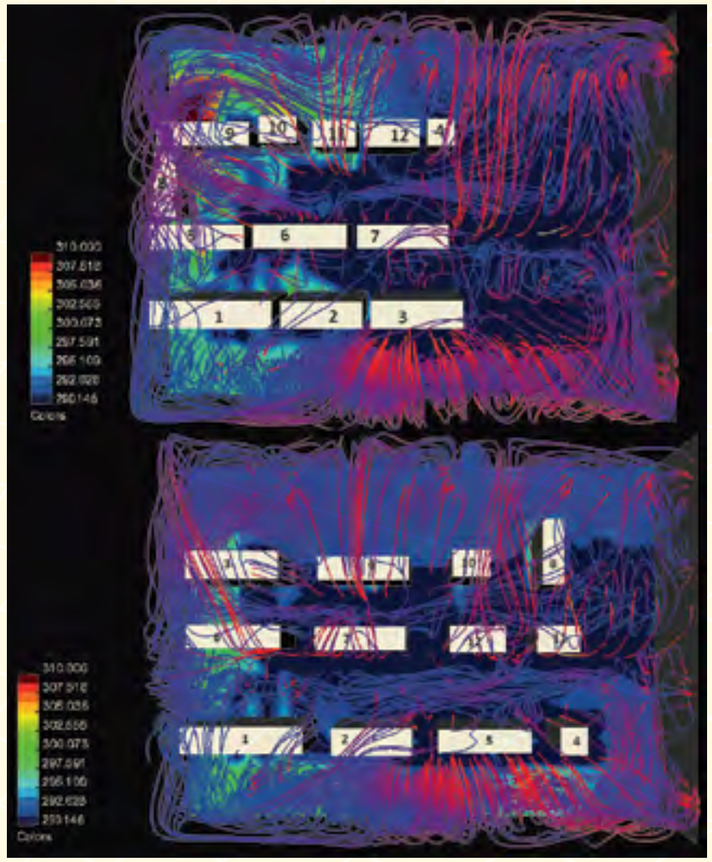


Figure 6: Rack placement with hotspot and optimized

placement lead to a more homogeneous temperature distribution in the room, especially for rack 9 the previous existing hot spot in the upper left corner can be solved (the temperatures are shown in Kelvin), the temperatures at the outlets are reduced by up to 30°K. Since the allowed maximum for the operational temperature is still the same, is might be possible to increase the temperature for the room inlets which will lead to a lower energy consumption for the cooling infrastructures since the difference between outside environmental air and coolant will be increased and in some cases even free cooling becomes possible with dramatically improved power efficiency. Based on Fig. 7 showing the temperature and speed of airflows recirculation can clearly be identified and countermeasures like building walls or changing the placement can be performed to prevent further recirculation and therefore reduced heat transfer. The negative effects of recirculation also demand an increase in air throughput to balance the heat transfer and so waste fan power and indirectly more cooling power for the additional fan power. Also higher temperatures may lead to increased failure rates if the temperatures reach the operational limits. With the simulation capabilities from CoolEmAll even the under-floor airflows can be checked and airflows can be assigned to each air-conditioner in the room to know the heat load distribution in advance before any installation is done. Fig. 8 show the scenario for three air conditioners from two perspectives where it can clearly be seen that the airflows are not even distributed and there are also some strong vortices which also means wasted fan power and maybe to less air in some areas of the room or racks. The SVD Toolkit can also be used to show the effect of increased room temperatures on the energy efficiency of the cooling infrastructure, in our test case it was possible to save approx 10% of overall power consumption by increasing the temperature above 22.5°C - just high enough to allow free cooling. In the project we also addressed the creation of blueprints of containerized modular data centers where depending on the load additional container can flexibly be

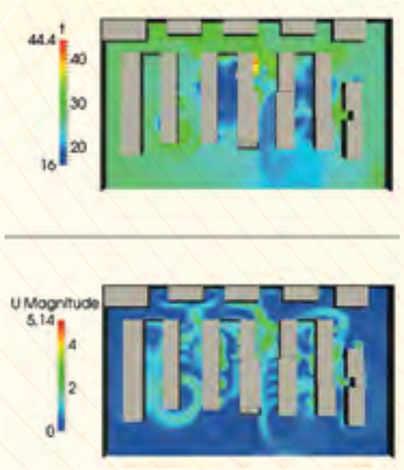


Figure 7: Recirculation within the room

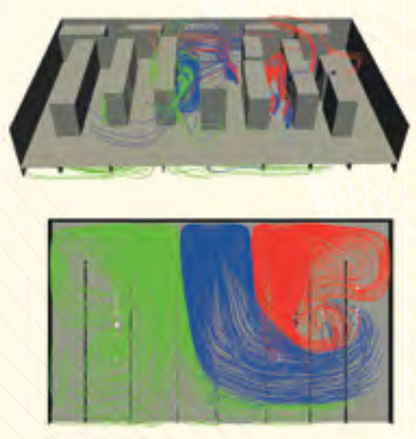


Figure 8: Uneven airflow distribution in the room and below the floor

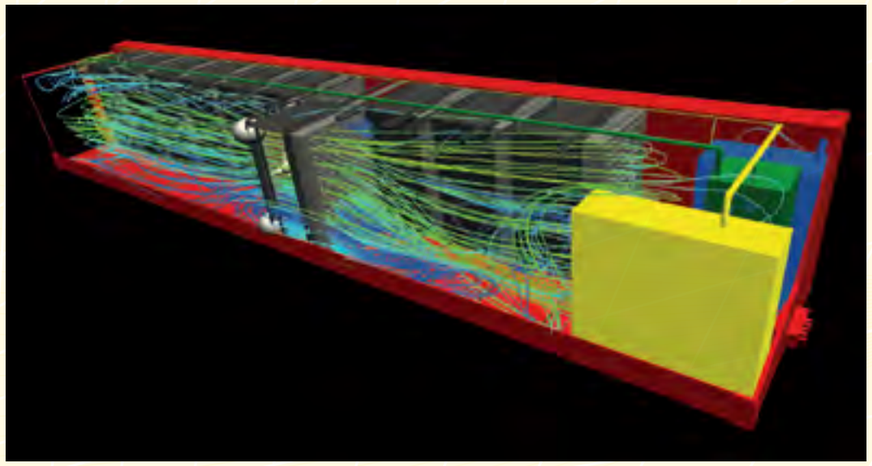


Figure 9: Containerized modular data centre

added as needed. An example of such a container equipped with RECS Boxes is given in Fig. 9 where airflows are illustrated.

The outcome of CoolEmAll allow to simulate data centres and optimize power consumptions before it is even build and whenever changes are required. The capabilities developed within the CoolEmAll project can be used to design data centres, describe all hardware devices, specify the expected applications and workload distribution and based on all this information simulate the operation of a data center to optimize the overall energy consumption. This leads to a more optimized rack placement, optimized temperatures and reduced need for fans and cooling power and may even be used to plan a better structured datacenter building.

CoolEmAll has been funded by the European Commission and started in October 2011 and has been ended in March this year.

Project Website:
http://www.coolemall.eu

Project Partners

- Poznan Supercomputing and Networking Center, PL
- High Performance Computing Centre University of Stuttgart, D
- Université Paul Sabatier - Institute for Research in Informatics of Toulouse, F
- christmann informationstechnik + medien GmbH & Co. KG, D
- 451 Research Ltd, UK
- The Catalonia Institute for Energy Research, E
- Atos, E

contact: Jochen Buchholz, buchholz@hlrs.de

- Jochen Buchholz
- Eugen Volk

University of Stuttgart, HLRS

MyThOS: Many Threads Operating System

For decades now, the computing industry has relied on increasing clock frequency and architectural changes to improve performance of CPUs. The situation has changed nowadays; small architectural improvements continue, but the power wall, the memory wall and thermal issues have made the approach of increasing clock frequency unfeasible. Semiconductor manufacturers continue developing the silicon manufacturing technology and doubling the number of transistor per unit area every few months, but rather than to improving the clock speed (because the physics do not allow it any more), they increase the number of processing units (i.e. cores) in the chip die.

This trend has not only been an uptake by HPC, but also by consumer computing and embedded systems, continuously increasing the core count as the main way to increase the (theoretical peak) performance. This creates a situation in which only embarrassing parallel applications can effectively improve the execution performance by taking advantage of the additional resources.

In order to maximize the theoretical performance of new system, developers must use an extremely high level of parallelism in their programs, rather than expect higher performance from the processor updates. But exploiting parallelism in any degree can be difficult: not only the data must be segmented, but the communication overhead can lead to reduced performance. Also, the amount of work of concurrent threads must be sufficient to compensate for the overhead of thread creation and management. Unrolling loops with short or little compute-intensive loop bodies can be too expensive and can reduce the performance of the program.

The adequate load per thread is dependent on several factors, the first is the code and data size. It also has to be taken into consideration the operations executed by the operating system to instantiate and configure the thread and the additional services needed to make the execution of the threads possible. Although in some operating system kernels implementations threads have fewer dependencies than processes, the minimum execution environment for threads is still too large and leads to cache misses, instantiation challenges and expensive context switches when a thread is rescheduled or invokes a system call.

This overhead is not just a problem that the programmer should cater for, but it is also a primary obstacle that keeps parallel applications from scaling better because it prevents the effective use of concurrency. Several studies [OEC11][TOL11] have noted that the current architecture of operating systems generally have a negative impact on scalability and performance. Parallel programming is a challenge for developers, and current approaches try to increase scalability by improving the data segmentation, producing better adapted thread bodies and reducing data dependencies. However, the operating system is part of the core of the problem, and has not yet been correctly addressed.

To correctly address this issue the architecture of operating systems for highly scalable applications has to be redesigned to eliminate secondary functionality and reduce as much as possible the overhead. The monolithic organization of current operating systems created dependencies between features that are not relevant for the execution of individual threads but affect the use of memory and computational resources. Instantiation of new threads require several 10,000s of compute cycles [OEC11][BAU09][BOY10][LIN00], and even if a thread pool can reduce this time, it also reduces the degree of the dynamicity that can realize the program. We expect the MyThOS operating system kernel design to be a

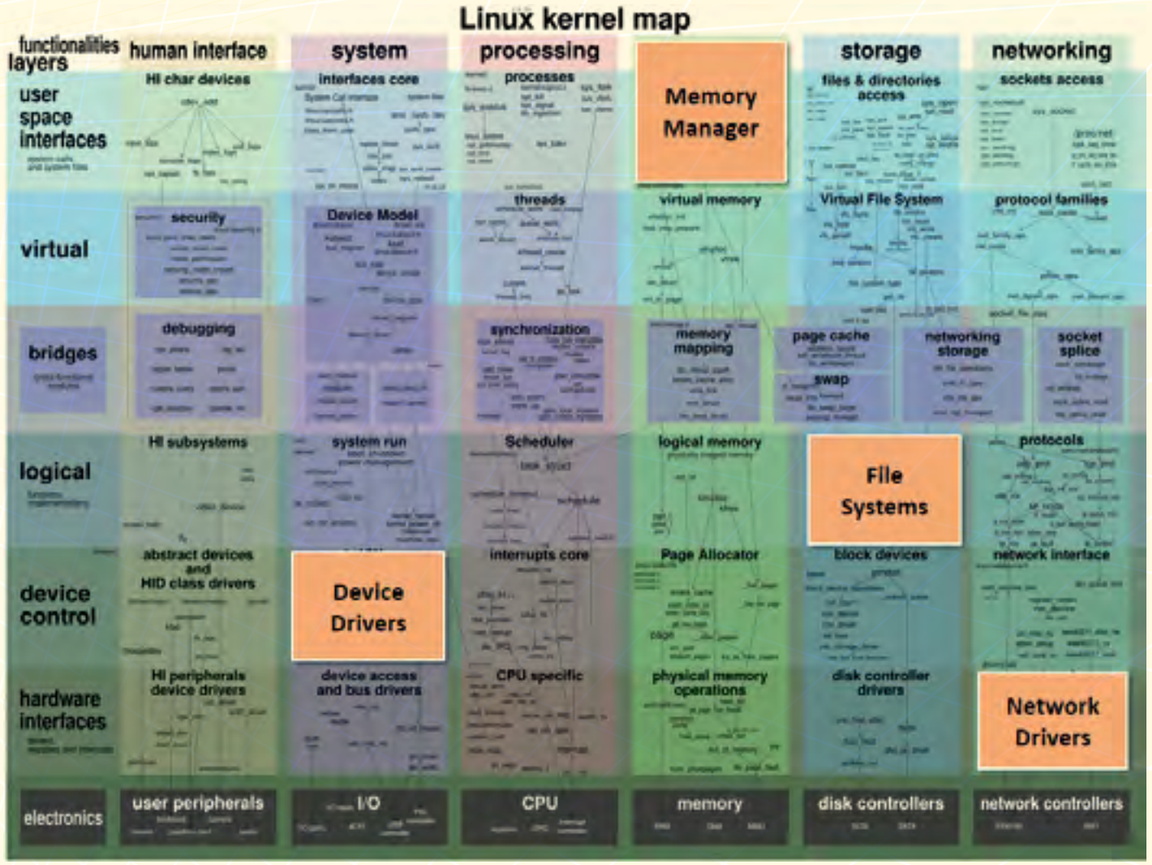


Figure 1: Linux Kernel Map, for the execution of a thread only a few modules are needed. (C) 2007-2010, http://www.makelinux.net/kernel_map

suitable solution for increasing the scalability of applications in multicore processors and multiprocessor nodes.

The MyThOS Project

The MyThOS project addresses a specific problem area that exists in both industry and academy and that is becoming increasingly relevant with the paradigm shift in processors' architecture. Even if the project is technologically innovative it is, at the same time, looking for direct applicability.

Current operating systems are designed to support a variety of architectures and application scenarios, thus not supporting the execution of specific use cases optimally. MyThOS aims to execute each application thread with just the necessary support for its execution. This implies that some aspects, such as security, time sharing, interactivity, etc., can be severely limited as the application will be given more control over the hardware. This is not a problem as for the project, compatibility with legacy software is secondary to performance and scalability, and, however, this does not contradict the requirements of the scientific, multimedia and compute intensive applications that MyThOS is targeting.

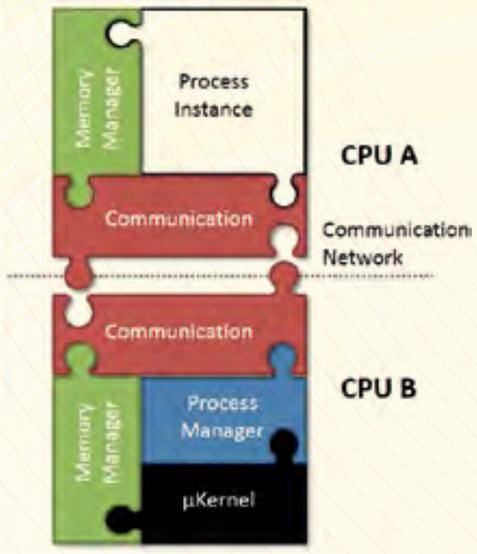


Figure 2: Different instances of the operating system on two CPUs

The project's aim is to reorganize the operating system architecture so that it is exclusively dedicated to tasks related to computing, specially multithreading, with some extra functionality necessary for creating the execution environment, communication and handling other necessary services for the system. These non-core functionality of the operating system are loaded accordingly to the needs of the system or the application at each moment. A lower operating system overhead means that threads with smaller bodies can be created, allowing for better exploitation of concurrency, increasing the scalability and the use of the available computing resources. All this together increases the application performance.

Technical Principles

Classic operating systems are based on a monolithic architecture, which does not allow for fast and dynamic use of parallelism due to the strong linkage of its functions and data structures. To solve this issue MyThOS will use a modular approach, taking concepts from the microkernel architecture paradigm, which have the implication that the computing resources will only be minimally loaded by the operating system, leaving more free resources to the program itself as in each computing unit there will be deployed only the minimal functionality for it to work. This will also reduce the operating system overhead by having less expensive system calls and improving the predictability of the system, which ultimately increases the performance of synchronous systems.

This modular approach also allows operating system functionality to be replaceable and distributed across the system as the functionality will be implemented in modules that use message passing as the means for communication, allowing programs to interchange data between threads instantiated on different system architectures, supporting communication across different system boundaries (i.e. between processor cores or between processor nodes, etc.).

To achieve maximum flexibility for the execution, applications running on MyThOS will be abstracted from the peculiarities of the different hardware memory address spaces and communication protocols so that a single uniform memory space and messaging system is presented. For an application there is no difference accessing local memory or memory that is not physically joined as the appropriate management is done by the operating system as the different system calls and memory accesses are handled transparently across all boundaries of the system's architecture. Similarly, system calls can be executed locally or remotely depending on where the necessary functionality and resources are available, and they interact with each other without the need to explicitly use the appropriate communication protocol for each case.

The core of the operating system (those components that involve messaging, memory management and thread instantiation) is loaded by default at system boot. This is the essential functionality to initialize processing units, to interact with each other, to load extra functionality and to start the execution of the application. As the application runs the distribution of the operating system and its modules providing the extra functionality will change accordingly with the needs of the application and the status of the system. Each process and each thread is allocated with the system components that provide the necessary functionality for an efficient execution of its task. For a better use

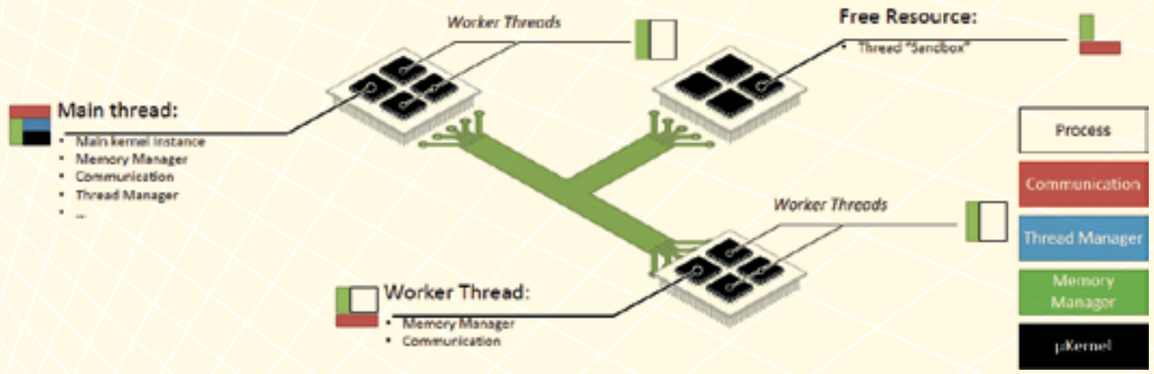


Figure 2: Architecture of distributed operating system with process and threads

of the resources, different variants of the same system component will be provided, optimized for different processors or system architectures, allowing to tune the system for fast deployment and management of threads on specific hardware. A single thread is allocated per core in low-weight "sandboxes", which can be quickly instantiated, configured and later destroyed with the aim to achieve a hierarchical multithreaded organization of the execution of the application that allows for dynamic management, changing its execution structure (the number and distribution of threads) according to the requirements of the application and resources available in the system at each moment.

To make this execution model possible, the individual operating system components must be orchestrated and coordinated with each other to make an efficient use of the resources and lower the communication overhead. MyThOS will stand on the research results of S(o)OS [SCH09a][SCH09b] and BarrelFish [BAU09] projects which clearly show that a distributed organization can significantly improve scalability and performance.

The project will be developed using Intel MIC, the Knight's Corner processor family, as its high CPU core count, internal interconnection models and memory organization, allow to experiment with a wide range of future hardware scenarios. MyThOS will give support for writing applications using common programming languages such as C and Fortran, which are widely used in industry as well as academia. The developments will be validated within the project on challenging use case scenarios from molecular dynamics, fluid dynamics and distributed computation of multimedia data. The scope of scenarios addressed in MyThOS however goes beyond the use cases MyThOS is well suited for analyzing big data and for highly dynamical scenarios.

To be able to use the capabilities proposed by MyThOS it is necessary to break compatibility with POSIX system calls, used in Unix-like operating systems, as otherwise we would face the limitations of application design and scalability that current systems have.

During the run time of the project it will not be possible, neither sensible, to develop the complete OS functionality, however, a prototype version of the operating system and its programming model will be developed, as well as documentation on how to extend it and how to design applications that can use the system capabilities correctly will be provided. Thanks to the modular design and implementation of MyThOS it will be easily further developed and extended in the future with new capabilities and support for future hardware architectures.

Who is MyThOS?

MyThOS is funded by the BMBF ("Bundesministerium für Bildung und Forschung", German Federal Ministry of Education and Research) and started in October 2013. The project partners include: University of Ulm; Brandenburg University of Technology; HLRS, University of Stuttgart; University of Siegen; Bell Labs, Alcate-Lucent AG.

References

[1] **Oechslein, B., Schedel, J., Kleinöder, J., Bauer, L., Henkel, J., Lohmann, D., Schröder-Preikschat, W.**
OEC11: OctoPOS, A Parallel Operating System for Invasive Computing, 2011.

[2] **van Tol, M.W.**
TOL11: A Characterization of the SPARC T3-4 System, 2011.

[3] **Baumann, A., Barham, P., Dagand, P.-E., Harris, T., Isaacs, R., Peter, S., Roscoe, T., Schüpbach, A., Singhania, A.**
BAU09: The multikernel, a new OS architecture for scalable multicore systems, 2009.

[4] **Boyd-Wickizer, S., Clements, A.T., Mao, Y., Pesterev, A., Frans Kaashoek, M., Morris, R., Zeldovich, N.**
BOY10: An analysis of Linux scalability to many cores, 2010.

[5] **Ling, Y., Mullen, T., Lin, X.**
LIN00: Analysis of Optimal Thread Pool Size, 2000.

[6] **Schubert, L., Kipp, A.**
SCH09a: Principles of Service Oriented Operating Systems, 2009.

[7] **Schubert, L., Kipp, A., Wesner, S.**
SCH09b: Above the Clouds: From Grids to Service- oriented Operating Systems, 2009.

contact: Colin W. Glass, glass@hlrs.de

- Colin Class
- Daniel Rubio Bonilla

University of Stuttgart, HLRS

Projects

HPC for Climate-friendly Power Generation

The joint research project CEC [1] funded by the Federal Ministry of Economics and Technology (BMWi) and the Siemens AG was launched in January 2013 with an initial three year funding phase. Besides Siemens – the project coordinator - and JSC the consortium includes seven German research institutes in the fields of power plant technology and material science. Aim of the project is to develop new combustion technologies for climate-friendly power generation. The focus is on modern gas turbine technology which play an important role in the present transformation of the energy system towards a sustainable system based on renewable energy sources. The validation of the gas turbine combustion technologies will be carried out in a new test center called "Clean Energy Center" (CEC) which is currently being built by Siemens in Ludwigsfelde near Berlin.

While testing is a key step in the validation of new gas turbine combustion systems, numerical methods can provide significantly more detailed insight into the behavior of gas turbine combustion systems. The detailed insight is important since it enables engineers to derive new design hypotheses needed to achieve higher thermodynamic efficiencies and reduced pollutant emissions. A key design difficulty in the development of modern low NOx gas turbine combustion systems are thermoacoustic instabilities. Thermoacoustic instabilities can result in pressure oscillations which require the immediate shut down of a power plant to avoid mechanical failure. Avoiding these events is a high priority. Developing a modeling capability for thermoacoustic instabilities is challenging. The phenomenon involves unsteady aerodynamics, chemical reaction and acoustics resulting in a stiff multi-physics problem.

Computational fluid dynamics on massively parallel computers was found to be a suitable approach to predict thermoacoustic instabilities for single burner systems. To resolve all relevant effects simulations consuming approx. 200,000 core-hours are needed. While the behavior of a single burner is similar to the behavior of a can combustor based gas turbine, annular combustor based gas turbines require the simulation of the complete combustion system comprising up to 24 burners (see Fig. 1).

The resulting computational requirements are difficult to satisfy on PC technology based clusters – larger scale systems are needed.

Figure 1: Gas turbine with can combustion system (lhs) vs. annular combustion system (rhs)

OpenFOAM for JUQUEEN

Thus one of the aims of the CEC project is the provision of CFD software which makes efficient usage of massively parallel HPC architectures. Within the scope of this project the open source software package OpenFOAM [2] will be used.

Before the start of the CEC project OpenFOAM has already been used by Siemens for production runs on JSCs HPC cluster JUROPA. There simulations of a single combustor tube are done applying typically a few thousand cores / several hundred nodes. Within this project it is planned to compute the behavior of a complete annular combustion chamber which will lead to models where the size (number of CFD cells) is increased by at least a factor of 10. In order to keep the turnaround times acceptable it is necessary to increase the number of cores used for the calculations by a similar factor.

Thus as a first step OpenFOAM was implemented on JSCs Blue Gene/Q system JUQUEEN which is more suitable for simulations running on tens of thousands of cores. This was not straightforward since the dynamical linking strategy applied by OpenFOAM is not the first choice for highly scalable architectures like JUQUEEN. The resulting OpenFOAM installation was successfully tested and made available for use by Siemens as well as other research groups with access to JUQUEEN.

Benchmarking of OpenFOAM on JUQUEEN with a model of a single combustion tube provided by Siemens showed that the single core simulation speed is as expected less than on JUROPA due to the slower clock speed of JUQUEEN's A2 processor. But with evolving models OpenFOAM on JUQUEEN will be the only option to handle frequent production runs of the required sizes.

At present work at JSC is under way to analyze the scaling behaviour of OpenFOAM using modern performance analysis tools like Scalasca [3, 4] which is a joint development of JSC and the German Research School for Simulation Sciences. The main goal of the JSC work package is to reduce performance and scaling bottlenecks in OpenFOAM to enable large simulations on JUQUEEN.

Project Partners

- Siemens AG, Energy Sector, Mühlheim an der Ruhr
- Institut für Verbrennungstechnik, DLR
- Institut für Verbrennung und Gasdynamik (IVG), Universität Duisburg-Essen
- Institut für Thermische Strömungsmaschinen (ITS), KIT
- Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH
- Institut für Prozess- und Anwendungstechnik Keramik, RWTH Aachen
- Institut für Strömungsmechanik und Technische Akustik (ISTA), TU Berlin
- Zentrum für Angewandte Raumfahrttechnik und Mikrogravitation, ZARM, Universität Bremen
- Department Werkstoffwissenschaften, FAU Erlangen-Nürnberg

References

[1] http://www.fz-juelich.de/ias/jsc/EN/Research/Projects/_projects/cec.html

[2] http://www.openfoam.org

[3] **Geimer, M., Wolf, F., Wylie, B.J.N., Ábrahám, E., Becker, D., Mohr, B.** The Scalasca performance toolset architecture, Concurrency and Computation: Practice and Experience, 22(6):702-719, April 2010.

[4] http://www.scalasca.org

contact: Johannes Grotendorst, j.grotendorst@fz-juelich.de

- Christian Beck[1]
- Johannes Grotendorst[2]
- Bernd Körfgen[2]

[1] Siemens AG, Energy Sector, Mülheim a. d. Ruhr

[2] Jülich Supercomputing Centre (JSC)

Going DEEP-ER to Exascale

October 1, 2013 marked the start of the EU-funded project "DEEP-ER" (DEEP–Extended Reach). As indicated by its name, DEEP-ER aims for exztending the Cluster-Booster Architecture proposed by the currently running EU-project "DEEP" [1] with additional functionality. The DEEP-ER project will have a duration of three years and a total budget of more than 10 million Euro. Its coordinator, the Jülich Supercomputing Centre, leads a consortium of 14 partners from 7 different European countries.

The heterogeneous Cluster-Booster Architecture implemented by DEEP consists of two parts: a Cluster based on multi-core CPUs with InfiniBand interconnect, and a Booster of many-core processors connected by the EXTOLL network. DEEP-ER (see Fig. 1) will simplify this concept by unifying the interconnect merging Cluster and Booster Nodes into the same network. The DEEP-ER prototype will explore new memory technologies and concepts like non-volatile memory (NVM) and network attached memory (NAM). Additionally, with respect to DEEP the processor technology will be updated: next generation Xeon processors will be used in the Cluster and the KNL generation of Xeon Phi will populate the Booster.

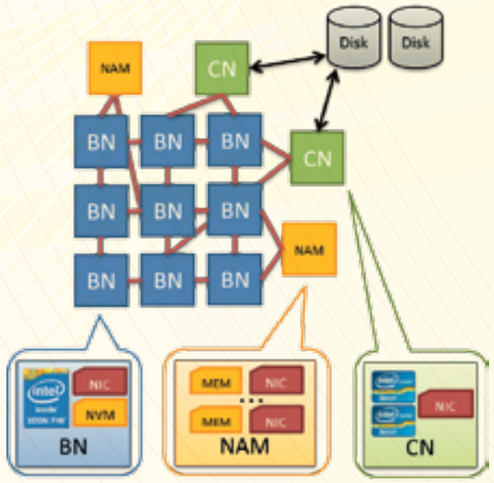


Figure 1: Cluster-Booster Architecture as implemented in DEEP-ER. (CN: Cluster Node; BN: Booster Node; NAM: Network Attached Memory; NVM: Non-Volatile Memory; MEM: Memory). In contrast to DEEP, here Cluster and Booster Nodes are attached to the same interconnect. Additional memory and storage is added at the node, the network, and system levels.

The DEEP-ER multi-level I/O infrastructure has been designed to support data-intensive applications and multi-level checkpointing/restart techniques. The project will develop an I/O software platform based on the Fraunhofer parallel file system (BeeGFS), the parallel I/O library SIONlib, and the I/O software package Exascale10. It aims to enable an efficient and transparent use of the underlying hardware and to provide all functionality required by applications for standard I/O and checkpointing.

DEEP-ER proposes an efficient and user-friendly resiliency concept combining user-level checkpoints with transparent task-based application restart. OmpSs is used to identify the application's individual tasks and their interdependencies. The OmpSs runtime will be extended in order to automatically restart tasks in the case of transient hardware failures. In combination with a multi-level user-based checkpoint infrastructure to recover from non-transient hardware-errors, applications will be able to cope with the higher failure rates expected in Exascale systems.

DEEP-ER's I/O and resiliency concepts will be evaluated using seven HPC applications from fields that have proven the need for Exascale resources.

Figure 2: Participants of the DEEP-ER kickoff meeting.

Acknowledgements

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 610476 (Project "DEEP-ER").

References

[1] **Suarez, E., Eicker, N., Gürich, W.** "Dynamical Exascale Entry Platform: the DEEP Project", inSiDE Vol. 9 No.2, Autumn 2011, http://inside.hlrs.de/htm/Edition_02_11/article_12.html

contact: Estela Suarez, e.suarez@fz-juelich.de

- Estela Suarez
- Norbert Eicker

Jülich Supercomputing Centre (JSC)

Score-E – Scalable Tools for Energy Analysis and Tuning in HPC

For some time already, computing centers feel the severe financial impact of energy consumption of modern computing systems, especially in the area of High-Performance Computing (HPC). Today, the share of energy already accounts for a third of the total cost of ownership, see Fig. 1, and is continuously growing.

The main objective of the Score-E project, funded under the third "HPC software for scalable parallel computers" call of the Federal Ministry of Education and Research (BMBF), is to provide user-friendly analysis and optimization tools for the energy consumption of HPC applications. These tools (Scalasca [2], Vampir [3], Periscope [4], and TAU [5]) will enable software developers to investigate the energy consumption of their parallel programs in detail and to identify program parts with excessive energy demands, together with suggestions on how to make improvements and to evaluate them quantitatively. In addition, the project will develop models to describe not directly measurable energy-related aspects and a powerful visualization of the measurement results.

Measurement of energy consumption will take advantage of hardware counters like Intel's "Running Average Power Limit" (RAPL) counters introduced with the Sandy Bridge architecture, IBM's

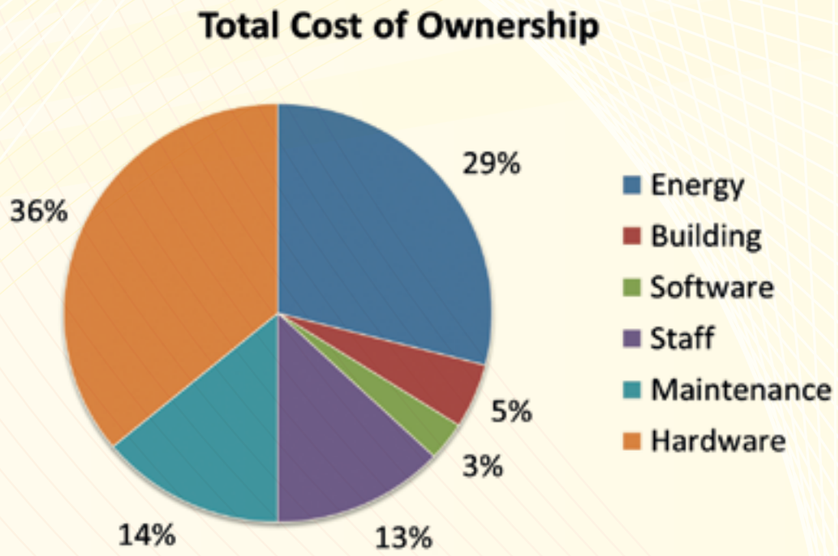


Figure 1: Total Cost of Ownership of a typical supercomputer according to a study by RWTH [1].

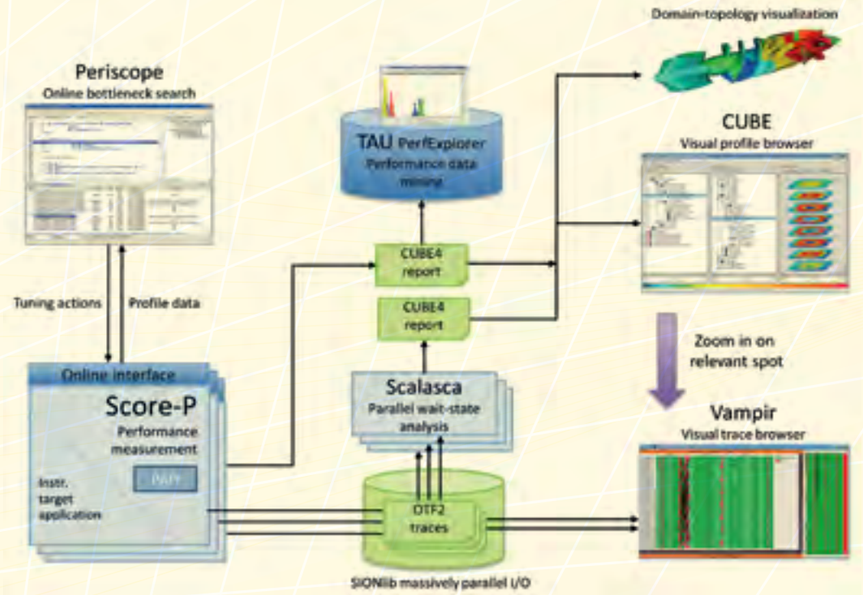


Figure 2: Interaction of the tools Persicope, Scalasca, Vampir, TAU and Scalasca's profile browser CUBE via the data exchange formats OTF2 for traces and CUBE4 for profiles.

Blue Gene/Q application programming interfaces to query power consumption on a node-board level, as well as system specific power measurement infrastructure on the node and rack level e.g., on SuperMUC. The new energy-related metrics can not only be visualized by the performance tools mentioned above, but also be used to trigger specific tuning actions during runtime by utilizing Score-P's online access interface.

With regards to visualization the Score-E project will investigate domain-topology visualizations as well as linked, multiple-view data presentation to gain deeper insight of the performance and energy behavior of complex applications.

At the same time, the project will further develop and maintain the community instrumentation and measurement system Score-P, see Fig. 2 [6], which forms the common base of all four tools mentioned above. To serve a broad user base, both, within the Gauss Alliance and beyond, most of the performance tools are released free of charge to the community under an open-source license. Only Vampir, due to its sophisticated user interface, is distributed commercially.

One particular task in the development will be the extension of the "Scalable I/O library for parallel access to task-local files" (SIONlib [7]) to support not only MPI plus basic OpenMP but arbitrary programming models and heterogeneity by providing a generic callback-based interface as well as a key-value scheme to handle a variable number of tasks per process. This extension widens the applicability of SIONlib to all

programming models currently in use in HPC.

The software products are accompanied by training and support offerings through the Virtual Institute–High Productivity Supercomputing (VI-HPS [8]), and will be maintained and adapted to emerging HPC architectures and programming paradigms beyond the lifetime of the Score-E project itself.

Besides the evident economic and environmental benefits in terms of energy, Score-E will also empower the optimized programs to unlock new scientific and commercial potentials.

The academic project partners in LMAC are the Jülich Supercomputing Centre, the German Research School for Simulation Sciences, RWTH Aachen University, TU Dresden and TU Munich. The industrial partner GNS mbH, a private company that specializes in services related to metal forming simulations, such as mesh generation for complex structures and finite element analyses, coordinates the project.

In addition, the University of Oregon, an associated partner, complements the Score-E objectives with corresponding extensions to the performance tool TAU. Further associated partners are Engys UG, who specializes in the application, support and development of Open Source Computational Fluid Dynamics (CFD) software and Munters Euroform which its expertise in engineering droplet separation systems for various industrial purposes.

For more information see:
http://www.vi-hps.org/projects/score-e
and http://www.score-p.org

- Christian Rössel [1]
- Bernd Mohr [1]
- Michael Gerndt [2]

[1] Jülich Supercomputing Centre (JSC)

[2] Technische Universität München

Acknowledgements

The Score-E project is funded by the German Federal Ministry of Research and Education (BMBF) under Grant No. 01IH13001.

References

[1] Bischof, C., an Mey, D., Iwainsky, C. Brainware for green HPC Computer Science–Research and Development, 27(4):227–233, 2012.

[2] http://scalasca.org/

[3] http://www.vampir.eu/

[4] http://www.lrr.in.tum.de/~periscop/

[5] https://www.cs.uoregon.edu/research/tau/home.php

[6] http://www.score-p.org

[7] http://www.fz-juelich.de/ias/jsc/EN/Expertise/Support/Software/SIONlib/_node.html

[8] http://www.vi-hps.org

Christian Rössel,
c.roessel@fz-juelich.de

UNICORE 7 Released

The UNICORE middleware suite is well-established as one of the major solutions for building federations and e-infrastructures, with a history going back to 1996 [1]. It is in worldwide use in HPC oriented infrastructures, for example PRACE, the US-project XSEDE [2] and in national grid initiatives such as PL-Grid. This spring, UNICORE 7 was released, which is the first major release since UNICORE 6.0 in August 2007.
It is no paradigm change as was the change from UNICORE 5 to UNICORE 6. Instead, UNICORE 7 is built on the same ideas and principles as UNICORE 6, and the two versions are compatible. We decided to make this a major release due to a number of improvements that serve to put the software on an updated technological basis.

As the most prominent change, we updated the internal web services stack to use the Apache CXF framework [3], which is the most advanced and mature Java services stack available today. This allows to build both WS/SOAP services as are currently used in UNICORE and RESTful services that will become more and more important in the future.

As the major new feature, we have added the possiblitity to deploy and run UNICORE in a way that end-users do not need certificates, using the Unity group management and federated identity solution [4]. Instead of using X.509 certificates to identify themselves, UNICORE clients request a signed Security Assertion Markup Language (SAML) [5] document from the Unity service, which is validated by the UNICORE services to assert the user's identity. Nevertheless, the strong client authentication based on X.509 certificates is still available and will continue to be supported in future releases.

Apart from the enhanced web services container and improved security stack, there are a number of other new features. For example, a new data-oriented processing feature allows to define data processing via user-defined rules. Jobs can be restarted easily, and data staging now supports wildcards.

Several changes have been made to improve the performance of UNICORE 7. For example, security sessions have been introduced to reduce the amount of XML data transferred between client and server, also reducing the CPU time required to process the XML messages. Several new batch operations have been added, for example allowing to delete multiple files or to check the status of many jobs using a single request/reply web service call. In data staging, the transfer of directories or multiple files has been optimized. Now, multiple files can be transferred in a single session, greatly using the overhead. This works especially well in conjunction with the UFTP high-performance data transfer protocol.

Together with the UNICORE 7.0 release, a first version of the new UNICORE Portal component was made available. This serves the increasing demand of users and infrastructure operators for a web-based access to UNICORE

resources. Compared to desktop clients, no installation on end user machines is required, and updates and fixes can be made available much easier. The UNICORE Portal was developed in a highly modular and customizable fashion, and allows deployers to disable unneeded functionality and to extend the portal with customized components, for example for user authentication. The underlying technology is Java, using Vaadin [6], an open-source web framework which allows for a rich user experience comparable to desktop clients.

Special care has been taken to allow integrating the Portal with different user authentication methods. The user can authenticate via X.509 certificate in the browser, or via the Unity identity management service, which allows also username/password authentication. Other methods such as authentication via Kerberos are possible, too, since the security subsystem is extensible.

The Portal currently offers access to the most important UNICORE features, which is job submission and monitoring, data access and transfer, and access to a subset of workflow features.

The screenshot (Fig. 1) shows the data browser view, which is a two-window component inspired by tools such as Norton Commander. It allows to manage data on both UNICORE storages and the end user's local filesystem. Data can be uploaded to, downloaded from or transferred between UNICORE servers in a very simple way. The user's local file system is made accessible using a Java applet.

UNICORE 7 is a major step forward in many respects, and we expect it will be quickly and widely deployed by our users in PRACE, XSEDE, national infrastructures such as PL-Grid, and others.

A particularly interesting new deployment of UNICORE is under development in the FET-Flagship „Human Brain Project“, where UNICORE will form the basis for the HPC platform, which will combine High-Performance Computing and scalable storage into a future exascale platform for simulating the human brain. The Human Brain project is expected to bring new requirements, such as interactive supercomputing, large user federations, and others. UNICORE is well-placed to take on these requirements. A strong focus of further development will be put on the UNICORE portal, where we plan to add support for some of the more advanced UNICORE features such as metadata management. Customization of the portal for specific use cases and applications will play a big role, e.g. through the development of an application integration layer. Last not least, we will work on adding social features and “teamwork” functions, leveraging the group membership management capabilities of the Unity system to allow simple sharing of task definitions or data files, as well as receiving notifications about activities of group members.

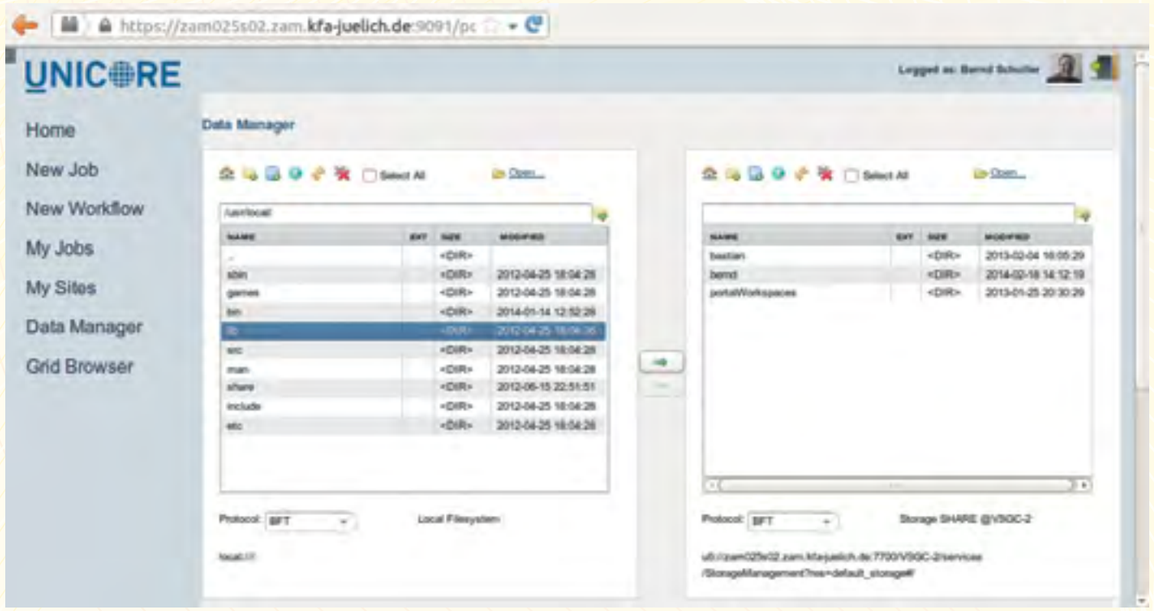


Figure 1: Screenshot of the data browser view in the UNICORE Portal

References

[1] **Streit, A. et.al.** annals of telecommunications 65, pp. 757–762, 2010

[2] UNICORE in XSEDE– https://www.xsede.org/software/unicore

[3] Apache CXF–http://cxf.apache.org

[4] Unity identity management solution– http://www.unity-idm.eu

[5] SAML–http://en.wikipedia.org/wiki/Security_Assertion_Markup_Language

[6] Vaadin web application framework– http://vaadin.com

contact: Bernd Schuller, b.schuller@fz-juelich.de

- Bernd Schuller
- Daniel Mallmann

Jülich Supercomputing Centre (JSC)

Project Mr. SymBioMath

The project High Performance, Cloud and Symbolic Computing in Big-Data Problems applied to Mathematical Modelling of Comparative Genomics (Mr.SymBioMath; http://www.mrsymbiomath.eu) represents an interdisciplinary effort focused on big data processing within the application domains of bioinformatics and biomedicine. The project partners cover a wide range of biological, medical and technological expertise. The University of Malaga (UMA), Spain acts as project coordinator and provides expertise in bioinformatics, comparative genomics as well as High Performance Computing and Cloud Computing. RISC Software GmbH, Austria is a software company with a special expertise on advanced computing technologies like cloud and Grid Computing. Besides contributing to the software development effort within the project, RISC Software GmbH provides the Cloud Computing infrastructure for the project. The Johannes Kepler University Linz, Austria contributes specific expertise on bioinformatics algorithms to the project, while Integromics – as a commercial company – will specifically address the commercialization of the project results, while also contributing to the core software development efforts. The project partner Servicio Andaluz de Salud–Hospital Carlos Haya, Spain is in charge of the biomedical use cases, which will be addressed by the project. Specifically this refers to discovering links between allergies and the relevant parts of the genomes of patients. The Leibniz Supercomputing Centre (LRZ), Germany contributes its expertise on visualization for the representation of the results of the bioinformatics and biomedical computations on different types of devices ranging from mobile devices to Virtual Reality devices.

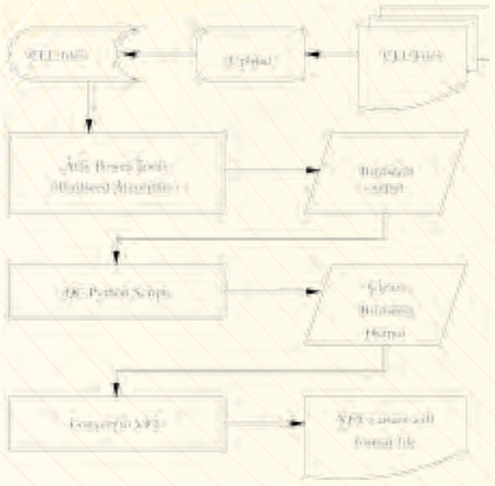


Figure 1: Biomedical Workflow

Infrastructure

Since the project is clearly focused on using Cloud Computing as its resource provision methodology a community cloud installation provided by project partner RISC Software GmbH is used for development and testing purposes. This cloud installation also gives the project consortium the opportunity to try out different cloud technologies for supporting the bioinformatics and biomedical applications within the project. As a Cloud Computing middleware OpenStack (http://www.openstack.org) is being used, which provides the users with access to different Cloud Computing services like execution of virtual instances, access to volumes acting as additional virtual hard disks as well as access to data containers where files can be stored to be exchanged between instances. This functionality can be accessed either through a web interface or through web-service calls, which enables the use by external programs like workflow engines. The role of workflow engines is crucial for the Mr. SymBioMath project since the bioinformatics and biomedical applications consist of a set of data transformation modules which can be flexibly linked to each other to achieve the desired overall functionality.

Bioinformatics and Biomedical Algorithms

Considering application domains the project is clearly focused on two related ones, bioinformatics and biomedicine. On the one hand the project works on genome comparison on the genome level and related applications, while on the other hand the biomedical aspects of the program are driven by the project partner Hospital Carlos Haya, which is specifically interested in the relation between genome variations and allergic reactions of patients. Both application areas require the processing of large genomic datasets and thus represent a good use case for the infrastructure's data handling capabilities. The specific bioinformatics and biomedical algorithms rely on the underlying infrastructure and provide their services to the higher layers of the hierarchy, represented by the end-user applications.

End-User Applications and Visualization

The main applications for end-users will be a visualization application focused on genome comparison and related functionality as well as a workflow composition application allowing a user to set up bioinformatics and biomedical workflows. The end-user applications will be designed to run on a range of different devices spanning from mobile devices like cell phones and tablets over desktop and laptop computers to Virtual eality devices like for example CAVEs (Cruz-Neira et al. 1992).

Workflows

The processing of workflows represents a core element of the software development within Mr. SymBioMath since they are applied both for the biomedical and bioinformatics domains. A first prototype implementation of a biomedical workflow as shown in Fig. 1 has been run on the Mr. SymBioMath cloud infrastructure by executing software modules consisting of preexisting C code and python scripts (Heinzlreiter et al., 2014). The workflow starts with CEL files containing single nucleotide polymorphisms (SNPs)–areas of variation within the genomes–from different human patients. The amount of data to be processed can range from a few megabytes to terabytes depending on the number of patients participating in a specific study. After the files have been uploaded the birdseed algorithm (Korn et al., 2008) is applied to them. After this step the algorithms output data has to be filtered to remove areas which are not varying across different patients and where the birdseed algo-

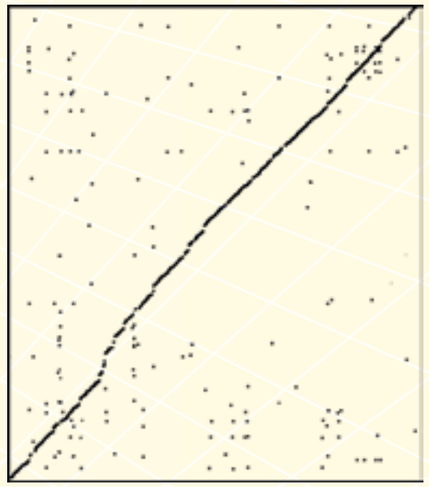


Figure 2: Example Dotplot

rithm did not produce appropriate results. To conclude this workflow, the filtered result has to be converted to the VCF file format, which represents a standard file format for genomic data.

The workflow was instantiated by executing preexisting software modules – a C code for the birdseed algorithm and Python scripts for the other steps – being distributed across multiple different cloud instances. The data exchange between the modules was realized by storing files into OpenStack containers. To enable accessing the containers from the running cloud instances the containers have been mounted into the file systems of the instances. This was realized by the cloudfuse-mount daemon which accesses the OpenStack containers through web service calls and makes it locally accessible in a way comparable to Network File System (NFS) mounts. With this infrastructure in place the workflow was distributed across different instances according to the computational load being generated by the different modules. During test runs, the last step of the workflow – the conversion to the VCF format – has been identified as the most time-consuming step. Typically the workflow gets executed with several CEL files as input which allows for the parallelization of the VCF-conversion step: The filtered output of the birdseed algorithm also consists of several files, which can be processed independently. Therefore the execution of the last step of the workflow was performed by running it on several instances in parallel.

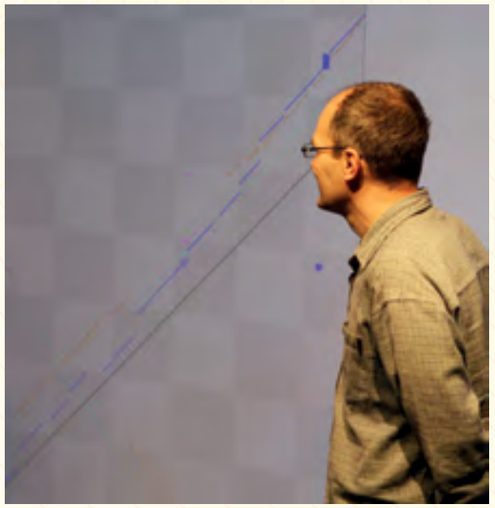


Figure 3: Genome Visualization

While the workflow described here is a very simple first step illustrating the viability of workflow execution on the Mr. SymBioMath cloud infrastructure, this setup is by no way user-friendly and needs to be improved and extended heavily to be usable for its actually intended users like bioinformaticians and clinicians. To make the execution of workflows more user-friendly and flexible a workflow engine like Galaxy (http://galaxyproject.org/) will be used. The workflow engine will provide a user-friendly way of setting up workflows and enacting them. Besides that, the storage of workflows will enable different groups of researchers to repeat experiments thus supporting scientific verification of published results.

A different example workflow which will be executed on the Mr. SymBioMath infrastructure focuses on bioinformatics and there more specifically on genome comparison. The overall idea of this workflow is given by the comparison of two genome sequences and the graphical representation of the results in a dotplot with an example shown in Fig. 2. A dotplot shows the similarities between two sequences by representing one genome sequence on the x-axis and the other one on the y-axis. If a similarity occurs between the two sequences it is marked with a dot at the corresponding (x, y) coordinates. The computational steps involved in this workflow will also be executed in a distributed way across multiple instances of the cloud infrastructure, but as a next step this workflow will act as test case for the invocation through a workflow engine like Galaxy.

Client Software

Within the bioinformatics domain web-services are a common methodology to expose computational services to external users. Therefore project partner UMA has developed the web-service client software jORCA (Martin-Requena at al., 2010) which can be used for calling web services in a unified way. It will represent a core tool for invoking Mr. SymBioMath workflows through web-service interfaces. Within the project a GlobusOnline (GO) plugin for jORCA has been developed, which enables the use of the GO services directly from the jORCA client software.

After the genome comparison has been executed, its result should typically be viewed as a dotplot. An OpenGL-based application facilitating a flexible visualization of genome comparisons is currently being developed at LRZ and will act as a visualization and partly post-processing tool for the results of the genome comparison. The core features of the visualization application are given by supporting a three-dimensional visualization of multiple genomes. An example screenshot from a Virtual Reality device is shown in Fig. 3.

Conclusion

The article provided a high-level overview of the project's aims and some initial developments. The core targets for the upcoming months are given by continued software development to extend the initial components adding new functionality and tightening their integration.

Acknowledgements

The Mr. SymBioMath project is running from February 2013 to January 2017 and is being funded by the European Union within the 7th framework programme for research as an Industry-Academia Partnerships and Pathways (IAPP) project under grant agreement number 324554.

References

[1] **Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R.V., Hart, J.C.**
The CAVE: Audio Visual Experience Automatic Virtual Environment, Communications of the ACM, Vol. 35, No. 6, 1992, pp. 64–72.

[2] **Heinzlreiter, P., Perkins, J.R., Torreno, O., Karlsson, J., Antonio, J.**
Andreas Mitterecker, Miguel Blanca, Oswaldo Trelles: A Cloud-based GWAS Analysis Pipeline for Clinical Researchers, in Proc. of the 4th International Conference on Cloud Computing and Services Science, 2014, to be published.

[3] **Korn, J., Kuruvilla, F., McCarroll, S., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P., Darvishi, K., Lee, C., Nizzari, M., Gabriel, S., Purcell, S., Daly, M., Altshuler, D.**
Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs, Nature Genetics, Vol. 10, No. 40, pp. 1253-1260, 2008.

[4] **Martin-Requena, V., Ríos, J., García, M., Ramírez, S., Trelles, O.**
jORCA: easily integrating bioinformatics Web Services, Bioinformatics, Vol. 26, No. 4, pp. 553-559, 2010.

contact: Paul Heinzlreiter, Paul.Heinzlreiter@lrz.de

- Paul Heinzlreiter
- Christoph Anthes
- Balazs Tukora

Leibniz Supercomputing Centre (LRZ)

Projects

ExaFSA–Exascale Simulation of Fluid-Structure-Acoustics Interactions

For a bit more than one year now, DFG's Priority Programme 1648 "Software for Exascale Computing" (SPPEXA) has been running, with 13 project consortia addressing the multi-faceted challenges of exascale computing. In each issue of inSiDE, one of those projects will present its agenda and progress, starting now with ExaFSA.

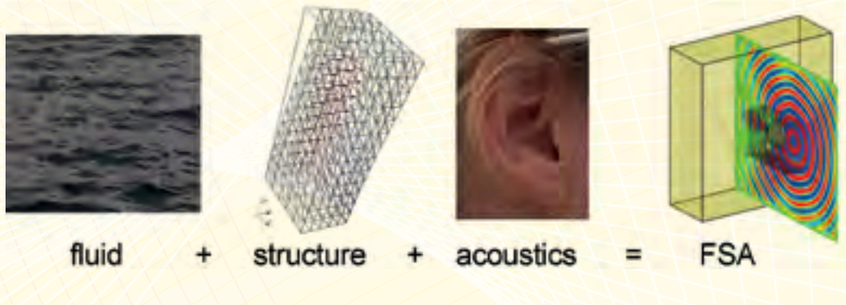


Figure 1: The project ExaFSA focuses on the three-field coupled simulation of fluid flow, structural dynamics, and acoustics.

The project team ExaFSA consisting of the groups of Miriam Mehl (Computer Science, Stuttgart), Hester Bijl (Aerospace Engineering, Delft), Thomas Ertl (Computer Science, Stuttgart), Sabine Roller (Computer Science, Siegen), and Dörte Sternel (Mechanical Engineering, Darmstadt) aims at tackling three of today's main challenges in numerical simulation: the increasing complexity of the underlying mathematical models, the massive parallelism of a supercomputer, and the huge amount of data produced as an output. The project focuses on interactions between fluids, structures, and acoustics (see Fig. 1) where application examples range from the sound design of ventilators, cars, aircrafts, and wind energy plants to simulations of the human voice. The huge amount of numbers produced as an output is to be translated into interpretable results using fast parallel in-situ visualization methods. As there are numerous highly sophisticated and reliable tools available for each of the three involved physical effects, a fast and flexible way to establish a fully coupled three-physics simulation environment is to combine such solvers.

We focus here on two particular aspects of our project work, both related to the coupling software gluing together the solvers, which is, besides efficient solvers for fluid flow, structural dynamics, and acoustics, a key component of the FSA simulation environment. We use the in-house coupling tool preCICE (see also [1] and Fig. 2), in which we recently implemented a new numerical method developed within our project that allows for an inter-solver parallelism without increasing the overall computational work compared to the standard traditional coupling of a flow and a structure solver that implies a mutual execution of the solvers which, due to non-equal workloads necessarily leads to a very inefficient usage of computational resources (see Fig. 3 and [2]).

The step from the mutual solver execution to the simultaneous execution implies a change of data dependencies as the structural solver cannot use the results of the current fluid solver step as an input anymore. If not handled carefully, this can lead to additional instabilities in particular for the challenging case of an incompressible fluid.

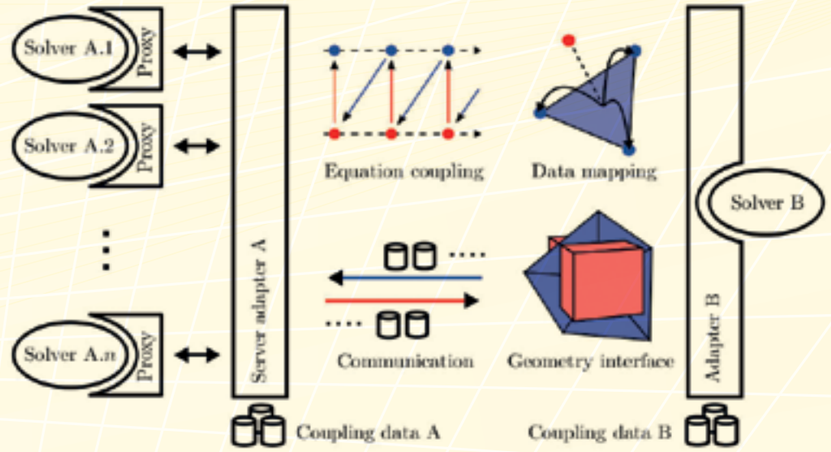


Figure 2: The coupling tool preCICE [1] provides the numerical equation coupling, data mapping, technical communication, and a geometry interface for coupled simulations. Parallel solvers such as solver A in the picture communicate over proxies with a server adapter running as a separate process that takes care of all numerical operations and the communication to the adapter of solver B which is a sequential solver in the shown example and runs in the same process as the solver B itself.

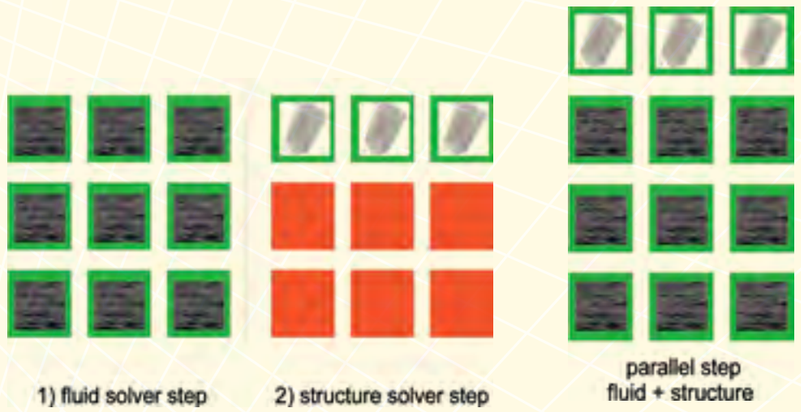
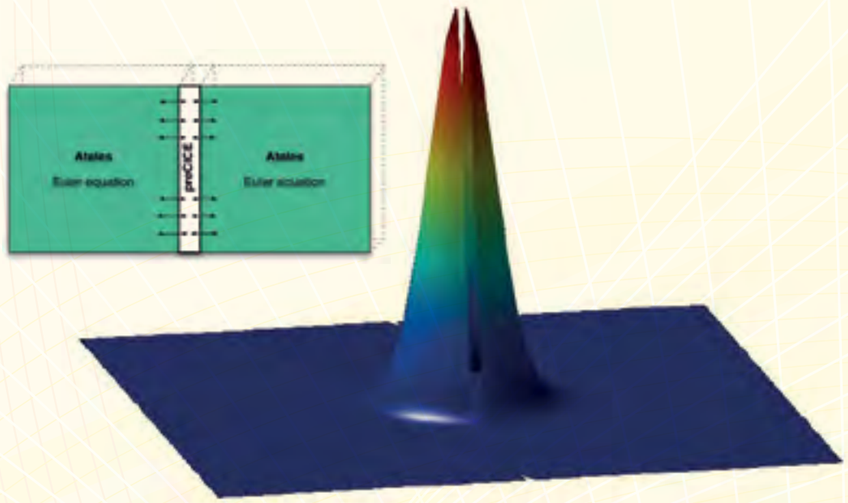
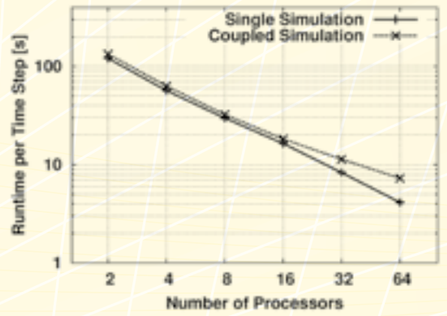


Figure 3: Imbalanced resource usage of the standard consecutive execution of fluid and structure solvers in a coupled fluid-structure simulation (a) versus our new balanced version where fluid and structure solver are executed simultaneously (b). Busy processors are marked green, idle waiting processors orange.



a)



b)

Figure 4: Simulation of a density pulse with Ateles. We compare runs using the usual MPI parallelization of Ateles with runs replacing the MPI communication at a vertical line through the domain by a communication via the coupling tool preCICE that we also use to couple fluid, structure and acoustics solvers. (a) Setup of the combination of Ateles with preCICE and simulation results showing the correctness at the coupling interface. (b) Strong scaling results for a total of 512 mesh elements, corresponding to 262144 degrees of freedom. The graph shows the bottleneck induced by sequential parts of preCICE which consequently have to be reduced to enable exascale FSA simulations. The simulations have been run at the MAC Cluster (http://www.mac.tum.de/wiki/index.php/MAC Cluster) in Munich. Full nodes, encompassing 16 physical processors, were reserved also for smaller runs, while runs with 32 or more processors used multiple nodes (e.g. 64 processors were run on 4 nodes). The preCICE server processes were always run on separated full nodes and are not taken into account for the scale up graph.

Our new coupling method overcomes such stability and convergence problems using a quasi-Newton approach where unknown internals of the solvers (which we assume to be black-box) are estimated from input-output observations and used in a suitable manner to stabilize and accelerate the solving of the system of equations related to each time step. Our results [2] show that, indeed, we can now solve the time step equation with the same number of fluid and structure solver steps as before with the mutual method, but almost twice as fast and without wasting energy leaving processors in a busy waiting state as we execute all components in parallel.

However, the coupling tool itself still can be a bottleneck as the following experiment shows: A simple spherical Gaussian pulse in density is transported with a constant fluid velocity through a domain governed by the nonlinear Euler equations. As this represents a pure transport phenomenon, the shape of this pulse should not vary over time. Thus, this setup is not only suitable for performance assessments, but also for validations of the results. We use the discontinuous Galerkin solver Ateles [3] with an 8th order scheme to solve this smooth flow problem. Due to the design of the DG-Method, relatively small amounts of data have to be communicated between high order elements in the mesh allowing the solver to scale down to single elements per process. The complete computational domain is then split for our experiment into two parts, where each part is computed by Ateles and the coupling of both is done via preCICE. The results (see Fig. 4) show that sequential parts of preCICE become a bottleneck at about a total of 32 processes, already, which leaves room for further improvements and innovations in the remaining project runtime.

Many aspects of FSA coupled simulations are typical for the larger class of multi-physics applications that are expected to profit from the project outcome as well. For more information about ExaFSA, further challenges in and contributions to efficient FSA simulations, see http://www.sppexa.de and http://ipvs.informatik.uni-stuttgart.de/SGS/EXAFSA/.

Acknowledgements

We thankfully acknowledge the financial support by the priority program 1648 – Software for Exascale Computing funded by the German Research Foundation (DFG).

References

[1] **Bungartz, H.-J., Benk, J., Gatzhammer, B., Mehl, M. and Neckel, T.**
Partitioned simulation of fluid-structure interaction on cartesian grids.
In: H.-J. Bungartz, M. Mehl, and M. Schäfer, editors, Fluid-Structure Interaction–Modelling, Simulation, Optimisation, Part II, volume 73 of LNCSE, pp. 255–284. Springer, Berlin, Heidelberg, October 2010.

[2] **Uekermann, B., Bungartz, H.-J., Gatzhammer, B. and Mehl, M.**
A parallel, black-box coupling for fluid-structure interaction.
In: Sergio Idelsohn, Manolis Papadrakakis, and Bernhard Schrefler, editors, Computational Methods for Coupled Problems in Science and Engineering, COUPLED PROBLEMS 2013, Stanta Eulalia, Ibiza, Spain, 2013. eBook (http://congress.cimne.com/coupled2013/proceedings/).

[3] **Zudrop, J., Klimach, H., Hasert, M., Masilamani, K. and Roller, S.**
A fully distributed cfd framework for massively parallel systems. In: Cray User Group, Stuttgart, 2012.

contact: Philipp Neumann,
neumanph@in.tum.de

- Hester Bijl [1]
- Thomas Ertl [2]
- Mehl, Miriam [1]
- Roller, Sabine [3]
- Dörte Sternel [4]

[1] Delft University of Technology

[2] University of Stuttgart (HLRS)

[3] Universität Siegen

[4] Technische Universität Darmstadt

The Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities (Leibniz-Rechenzentrum, LRZ) provides comprehensive services to scientific and academic communities by:

- giving general IT services to more than 100,000 university customers in Munich and for the Bavarian Academy of Sciences
- running and managing the powerful communication infrastructure of the Munich Scientific Network (MWN)
- acting as a competence centre for data communication networks
- being a centre for large-scale archiving and backup, and by
- providing High Performance Computing resources, training and support on the local, regional, national and international level.

Research in HPC is carried out in collaboration with the distributed, statewide Competence Network for Technical and Scientific High Performance Computing in Bavaria (KONWIHR).

Contact:
Leibniz Supercomputing Centre

Prof. Dr. Arndt Bode
Boltzmannstr. 1
85478 Garching near Munich
Germany

Phone +49-89-358-31-80 00
bode@lrz.de
www.lrz.de

Picture of the Petascale system SuperMUC at the Leibniz Supercomputing Centre.

Compute servers currently operated by LRZ are given in the following table

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
IBM System x "SuperMUC"	18 thin node islands IBM iDataPlex 512 nodes per island with 2 Intel Sandy Bridge EP processors each 147,456 cores 288 TByte main memory FDR 10 IB	3,185	Capability Computing	German universities and research institutes, PRACE projects (Tier-0 System)
	1 fat node island with 205 nodes IBM Bladecenter HX5 4 Intel Westmere EX processors each 52 TByte main memory QDR IB	78	Capability Computing	
	32 accelerated nodes with 2 Intel Ivy Bridge EP and 2 Intel Xeon Phi each 76 GByte main memory Dual-Rail FDR 14 IB	100	Prototype system	
Linux-Cluster	510 nodes with Intel Xeon EM64T/AMD Opteron 2-, 4-, 8-, 16-, 32-way 2,030 Cores 4.7 TByte	13.2	Capability Computing	Bavarian and Munich Universities, LCG Grid
SGI Altix ICE	64 nodes with Intel Nehalem EP 512 Cores 1.5 TByte memory	5.2	Capacity Computing	Bavarian Universities, PRACE
SGI Altix Ultraviolet	2 nodes with Intel Westmere EX 2,080 Cores 6.0 TByte memory	20.0	Capability Computing	Bavarian Universities, PRACE
Megware IB-Cluster "CoolMUC"	178 nodes with AMD Magny Cours 2,848 Cores 2.8 TByte memory	22.7	Capability Computing, PRACE prototype	Bavarian Universities, PRACE
MAC research cluster	64 Intel Westmere Cores, 528 Intel Sandy Bridge Cores, 1,248 AMD Bulldozer Cores, 8 NVIDIA GPGPU cards, 8 ATI/AMD GPGPU cards	40.5	Testing accelerated architectures and cooling technologies	Munich Centre of Advanced Computing (MAC), Computer Science TUM

A detailed description can be found on LRZ's web pages: www.lrz.de/services/compute

H L R S

First German National Center
Based on a long tradition in supercomputing at University of Stuttgart, HLRS (Höchstleistungsrechenzentrum Stuttgart) was founded in 1995 as the first German federal Centre for High Performance Computing. HLRS serves researchers at universities and research laboratories in Europe and Germany and their external and industrial partners with high-end computing power for engineering and scientific applications.

Service for Industry
Service provisioning for industry is done together with T-Systems, T-Systems sfr, and Porsche in the public-private joint venture hww (Höchstleistungsrechner für Wissenschaft und Wirtschaft). Through this co-operation industry always has acces to the most recent HPC technology.

Bundling Competencies
In order to bundle service resources in the state of Baden-Württemberg HLRS has teamed up with the Steinbuch Center for Computing of the Karlsruhe Institute of Technology. This collaboration has been implemented in the non-profit organization SICOS BW GmbH.

World Class Research
As one of the largest research centers for HPC HLRS takes a leading role in research. Participation in the German national initiative of excellence makes HLRS an outstanding place in the field.

Contact:
Höchstleistungsrechenzentrum
Stuttgart (HLRS)
Universität Stuttgart

Prof. Dr.-Ing. Dr. h.c. Dr. h.c.
Michael M. Resch
Nobelstraße 19
70569 Stuttgart
Germany

Phone +49-711-685-8 72 69
resch@hlrs.de / www.hlrs.de

View of the HLRS Cray XE6 "Hermit"

Compute servers currently operated by HLRS

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
Cray XE6 "Hermit" (Q4 2011)	3,552 dual socket nodes with 113,664 AMD Interlagos cores	1,045	Capability Computing	European and German Research Organizations and Industry
NEC Cluster (Laki, Laki2) heterogenous compunting platform of 2 independent clusters	23 TB memory 9988 cores 911 nodes	170	Laki: 120,5 TFlops Laki2: 47,2 TFlops	German Universities, Research Institutes and Industry

A detailed description can be found on HLRS's web pages: www.hlrs.de/systems

JÜLICH FORSCHUNGSZENTRUM

The Jülich Supercomputing Centre (JSC) at Forschungszentrum Jülich enables scientists and engineers to solve grand challenge problems of high complexity in science and engineering in collaborative infrastructures by means of supercomputing and Grid technologies.

Provision of supercomputer resources of the highest performance class for projects in science, research and industry in the fields of modeling and computer simulation including their methods. The selection of the projects is performed by international peer-review procedures implemented by the John von Neumann Institute for Computing (NIC), GCS, and PRACE.

Supercomputer-oriented research and development in selected fields of physics and other natural sciences by research groups and in technology, e.g. by doing co-design together with leading HPC companies.

Implementation of strategic support infrastructures including community-oriented simulation laboratories and cross-sectional teams, e.g. on mathematical methods and algorithms and parallel performance tools, enabling the effective usage of the supercomputer resources.

Higher education for master and doctoral students in cooperation e.g. with the German Research School for Simulation Sciences.

Contact:
Jülich Supercomputing Centre (JSC)
Forschungszentrum Jülich

Prof. Dr. Dr. Thomas Lippert
52425 Jülich
Germany
Phone +49-24 61-61-64 02
th.lippert@fz-juelich.de
www.fz-juelich.de/jsc

JSC's supercomputer "JUQUEEN", an IBM Blue Gene/Q system.

Compute servers currently operated by JSC

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
IBM Blue Gene/Q "JUQUEEN"	28 racks 28,672 nodes 458,752 processors IBM PowerPC® A2 448 Tbyte main memory	5,872	Capability Computing	European Universities and Research Institutes, PRACE
Intel Linux CLuster "JUROPA"	3,288 SMT nodes with 2 Intel Nehalem-EP quad-core 2.93 GHz processors each 26,304 cores 77 TByte memory	308	Capacity and Capability Computing	European Universities, Research Institutes and Industry, PRACE
Intel GPU Cluster "JUDGE"	206 nodes with 2 Intel Westmere 6-core 2.66 GHz processors each 412 graphic processors (NVIDIA Fermi) 20.0 TByte memory	240	Capacity and Capability Computing	selected HGF Projects
IBM Cell System "QPACE"	1,024 PowerXCell 8i processors 4 TByte memory	100	Capability Computing	QCD Applications SFB TR55, PRACE

Open Dialogue on Pre-Commercial Procurement of Innovative HPC Technology

Human Brain Project

The Human Brain Project (HBP) [1] is an ambitious, European-led scientific collaborative project, which is supported since October 2013 by the European Commission through its FET Flagship Initiative [2]. The HBP aims to gather all existing knowledge about the human brain, build multi-scale models of the brain that integrate this knowledge and use these models to simulate the brain on supercomputers.

Large-scale, memory-intensive brain simulations running on the future pre-exascale and exascale versions of the HBP Supercomputer will need to be interactively visualised and controlled by experimenters, locally and from remote locations. "Interactive Supercomputing" capabilities should allow the supercomputer to be used like a scientific instrument, enabling in silico experiments on virtual human brains. These requirements will affect the whole system design, including the hardware architecture, run-time system, mode of operation, resource management and many other aspects.

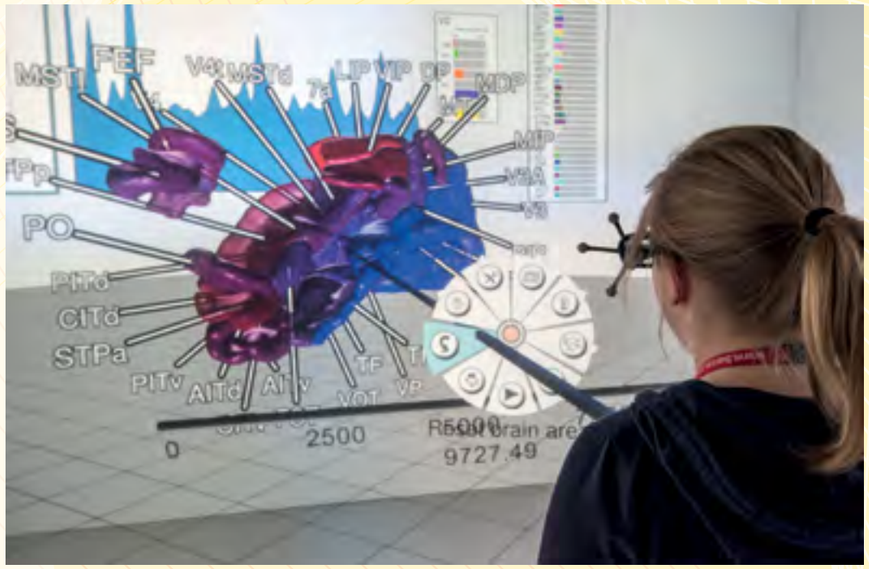


Figure 1: The image shows VisNEST [5], an interactive analysis tool for neural activity data, being used in the aixCAVE at RWTH Aachen University (Source: Virtual Reality Group, RWTH Aachen University).

While supercomputers with exascale capabilities will become available sooner or later without any HBP intervention, it is unlikely that these future systems will meet unique HBP requirements without additional research and development (R&D). JSC as the leader of the HBP's HPC Platform is therefore considering to carry out a Pre-Commercial Procurement (PCP) to drive R&D for innovative HPC technologies that meet the specific requirements of the HBP.

PCP is a relatively new model of public procurement, designed to procure R&D services, which is promoted by the European Commission (EC) [3]. It is organized as a competitive process comprising three phases: 1) solution exploration, leading to design concepts, 2) prototyping, 3) original development of limited volumes of first products. To select the supplier(s) best able to satisfy the goals of the PCP, the number of candidates is reduced at each phase after an evaluation of the respective results and the bids for the following phase. The risks and benefits of the PCP are shared between the procurer and the suppliers. For instance, the intellectual property rights (IPR) resulting from the PCP may remain with the suppliers if the procurer is granted a suitable license allowing the procurer to use the IPR.

An Open Dialogue event with interested potential suppliers was held by the HBP in Brussels on December 18, 2013 as part of a market exploration. Its goal was to present the HBP PCP framework and process, as well as a first version of the technical goals. The agenda included an overview of the HBP as a whole and the roadmap for building the HBP's HPC Platform. Discussion sessions offered potential suppliers opportunities to ask questions. The suppliers were encouraged to provide feedback, during the event and afterwards.

The event was attended by representatives of more than 25 different companies. While many of the questions raised during the event were aimed at getting a better understanding of the presented technical goals, the majority of the feedback concerned legal issues, in particular IPR regulations. The feedback received represents valuable information that is being taken into account in the drafting of the call for tender.

All slides presented during the meeting are available online [4].

Acknowledgements

The Human Brain Project receives funding from the European Union's 7th Framework Programme (FP7) under Grant Agreement no. 604102.

References

[1] https://www.humanbrainproject.eu/

[2] http://cordis.europa.eu/fp7/ict/programme/fet/flagship/

[3] http://cordis.europa.eu/fp7/ict/pcp/

[4] http://www.fz-juelich.de/SharedDocs/Termine/IAS/JSC/EN/events/2013/hbp-od-pcp-2013.html

[5] http://dx.doi.org/10.1109/BioVis.2013.6664348

contact: Boris Orth, b.orth@fz-juelich.de

- Dirk Pleiter
- Boris Orth

Jülich Supercomputing Centre (JSC)

Second JUQUEEN Porting and Tuning Workshop

Figure 1: Participants of the Second JUQUEEN Porting and Tuning Workshop (Source: Forschungszentrum Jülich GmbH).

Early last year the Jülich Supercomputing Centre (JSC) organized the First JUQUEEN Porting and Tuning Workshop. Encouraged by its success we decided to start a series of such events and followed it by the Second JUQUEEN Porting and Tuning Workshop from February 3 to 5. This year, as in the previous year, the goal of the course was to make the participants more familiar with the system installed at JSC and to provide tools and ideas to help with porting their codes, analyzing the performance, and in improving the efficiency. Apart from those topics, special attention was given to users from the field of Computational Fluid Dynamics (CFD) by providing dedicated talks and discussions.

Although the Blue Gene/Q installation at JSC with its 28 racks has been in, stalled for over a year now, this PRACE Advanced Training Centre (PATC) course again attracted 30 participants of which roughly half joined the CFD part of the workshop. The topics covered included in-depth talks on very specific hardware features of the Blue Gene/Q architecture like transactional memory and speculative execution alongside introductory talks to get the participants started. The latter consisted of best practices for programmers, overviews of available performance tools and debuggers, OpenMP, and parallel I/O. The main focus of the workshop was on hands-on sessions with the users' codes which were supervised by members of staff from JSC's Simulation Laboratories and cross-sectional teams (Application Optimization, Performance Analysis, Mathematical Methods and Algorithms) as well as from IBM.

Participants working on CFD had an additional series of talks on specialized CFD related subjects during the hands-on sessions. These talks were partly given by the participants themselves and focused on numerics, various CFD applications, and on experiences with CFD codes in HPC. In addition, every CFD group briefly presented the solver used in their application and discussed HPC related aspects. These talks greatly enhanced the communication and interaction between the research groups attending the workshop. This special interest group within the workshop was organized by the JARA-HPC simulation laboratory "Highly Scalable Fluids & Solids Engineering", which offers support on questions beyond the pure HPC domain, tailored to specific CFD software tools and applications.

Even though a workshop of only three days duration is too short to transfer all ideas from the talks to the users' codes, some immediate improvements were achieved. To provide an incentive to the participants, JSC awarded a prize for the best progress made during the course, which was decided via a vote on the last day.

The slides of the talks can be found on the web at

http://www.fz-juelich.de/ias/jsc/jqws14

contact: Dirk Brömmel, d.broemmel@fz-juelich.de

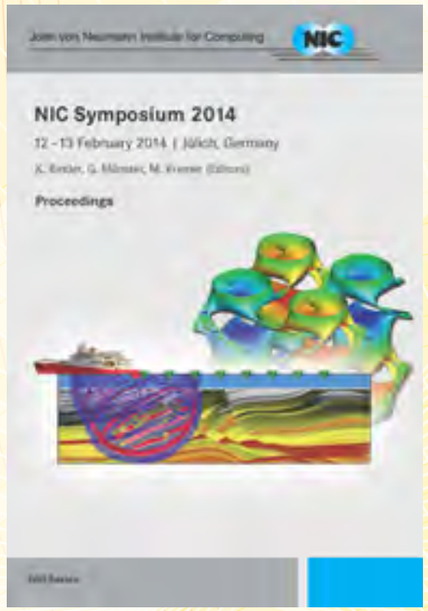
- Dirk Brömmel [1]
- Paolo Crosetto [1]
- Mike Nicolai [2]

[1] Jülich Supercomputing Centre (JSC)

[2] JARA-HPC

Activities

NIC Symposium 2014



The NIC symposium is held every two years to give an overview on activities and results obtained by research groups receiving computing time grants on the supercomputers in Jülich through the John von Neumann Institute for Computing (NIC). The 7th NIC Symposium took place at Forschungszentrum Jülich from February 12 to 13, 2014 and was attended by about 140 scientists.

The participants were welcomed by Prof. Achim Bachem (Board of Directors of Forschungszentrum Jülich), who portrayed current research in Jülich, and by Prof. Thomas Lippert (JSC), who presented the Jülich approach to pre-exascale supercomputing.

Prof. Kurt Kremer (MPI f. Polymer Science, Mainz) gave a historical account of soft matter computer simulations on the Jülich supercomputers provided through NIC and its predecessor, the Höchstleistungsrechenzentrum (HLRZ). He dedicated his talk to Prof. Kurt Binder on the occasion of his recent 70th birthday. Prof. Binder (University of Mainz) is one of the founding fathers of NIC/HLRZ [1] and currently serves as chairman of the NIC scientific council.

In the scientific programme, recent results in various fields of research, ranging from astrophysics to turbulence, were presented in 13 invited talks and in about 80 posters. Both, attendance and number of poster presentations were a new record high. Ample discussion sessions after the talks and the poster session gave the participants rich opportunities to exchange ideas and methods in an interdisciplinary setting. The detailed programme, talks, posters, proceedings, and pictures are available at http://www.fz-juelich.de/ias/jsc/nic-symposium/.

References

[1] see InSiDE Spring 2012, p.100.

contact: Walter Nadler,
w.nadler@fz-juelich.de

• Walter Nadler

Jülich
Supercomputing
Centre (JSC)

Figure 1: Participants of the NIC Symposium 2014 (Source: Forschungszentrum Jülich GmbH).

Opening Workshop of the Simulation Laboratory "Ab Initio"

Ab initio methods play a major role in the fields of chemistry, solid-state physics, nano-science and materials science. The impact of such methods has been steadily growing, currently generating computations which result in several thousands of scientific papers per year [1]. In order to keep up with this impressive output, it is crucial to preserve the performance and accuracy of ab initio simulations as the complexity of the physical systems under scrutiny is increased. Maintaining high performance and great accuracy poses significant challenges such as, for instance, simulating materials with interfaces, parallelizing and porting codes on new architectures, and converging the self-consistent field in Density Functional Theory (DFT) computations. In order to address these challenges, expertise in the fields of numerical mathematics, algorithmic development and hardware evolution is required.

There are three major factors hindering a straightforward achievement of performance and accuracy: 1) model diversity, 2) a need for scientific interdisciplinarity and 3) implementation heterogeneity. Model diversity emerges from the necessity of solving for a large spectrum of diverse scientific problems (e.g. band gaps, structural relaxation, magnetic properties, phonon transport, chemical bonds, crystallization, etc.) by simulating different aspects of a physical system. As a consequence ab initio methods have been realized in a rich variety of mathematical models, some examples of which are Density Functional Theory assorted discretizations, hybrid Quantum Mechanics/Molecular Mechanics methods and the Functional Renormalization Group approach, just to name a few. Driving the progress in all these methods requires an increasing cooperation within a multidisciplinary community which almost no working group worldwide can afford. Specifically there is a necessity to establish extensive collaborations between Physicists, Chemists, Computer Scientists and Applied Mathematicians.

Figure 1: Participants to the opening workshop of the ab initio Simulation Laboratory.

State-of-the-art codes are usually the result of a multi-person effort maintained over many years. The outcome is a large number of heterogeneous implementations, each one focusing on distinct physical aspects such as, for example, finite-size versus periodic systems or molecular versus solids compounds. Moreover, codes are written in a variety of styles and programming languages culminating in huge legacy codes which are difficult to maintain and port on new parallel architectures.

The Jülich Aachen Research Alliance–High-Performance Computing (JARA-HPC) [2] has established the Simulation Laboratory "ab initio methods in Chemistry and Physics" (SLai) [3] with the purpose of addressing all three main classes of issues. SLai is part of the strategic effort to improve the scientific collaboration between RWTH Aachen and the Forschungszentrum Jülich. Multidisciplinarity is deeply encoded in the DNA of the SimLab which connects the large community of ab initio application users with the code developers and the supercomputing support team. The mission of the Lab is to provide expertise in the field of ab initio simulations for physics, chemistry, nano-science and materials science with a special focus on High Performance Computing at the Jülich supercomputer facilities. It also acts as a high-level support structure in dedicated projects and hosts research projects dealing with fundamental aspects of code development, algorithmic optimization and performance improvement. Examples of current SLai activities are

- development of fast eigensolvers tailored to DFT methods;
- implementation of efficient tensor contraction kernels for Coupled Cluster methods;
- generation of improved data structures for FLAPW-based methods;
- development of a universal pre-conditioner for the efficient convergence of the charge density in DFT.

On November 8, 2013 SLai held its kick-off event at the Jülich Supercomputing Centre Rotunda Hall [4]. The workshop was designed to introduce the SimLab to the local condensed-matter physics and quantum chemistry communities, bringing together 50 scientists from institutes within JARA. By sharing their work and scientific expertise, the participants established a platform defining opportunities for mutual collaboration and prioritization of the SimLab's activities. Short informal talks were given by speakers from each of the main participating partner institutes and cross-sectional groups, followed by extensive constructive discussions between the speakers and the audience.

References

[1] Burke, K. Perspective on Density Functional Theory, J. Chem. Phys. 136, pp. 150901, AIP Publishing LLC 2012

[2] http://www.jara.org/en/research/jara-hpc/

[3] http://www.jara.org/hpc/slai

[4] http://www.jara.org/hpc/slai/kick-off

contact: Edoardo A. Di Napoli
e.di.napoli@fz-juelich.de

- Edoardo Di Napoli

JARA-HPC

Jülich Supercomputing Centre (JSC)

Activities

New Books in HPC

Tools for High Performance Computing 2012

Cheptsov, A., Brinkmann, S., Gracia, J., Resch, M., Nagel, W.E. (Eds.)
2013, XI, 162 p. 70 illus. in color.

- Written by the leading experts on the HPC tools market
- Describes best-practices of parallel tools usage
- Contains real-life practical examples.

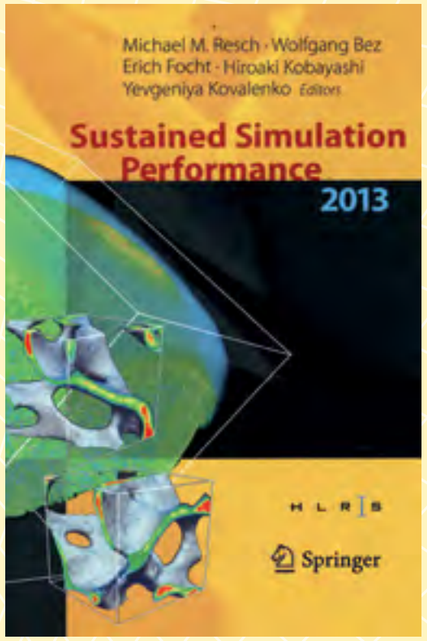
The latest advances in the High Performance Computing hardware have significantly raised the level of available compute performance. At the same time, the growing hardware capabilities of modern supercomputing architectures have caused an increasing complexity of the parallel application development. Despite numerous efforts to improve and simplify parallel programming, there is still a lot of manual debugging and tuning work required. This process is supported by special software tools, facilitating debugging, performance analysis, and optimization and thus making a major contribution to the development of robust and efficient parallel software. This book introduces a selection of the tools, which were presented and discussed at the 6th International Parallel Tools Workshop, held in Stuttgart, Germany, September 25 to 26, 2012.

High Performance Computing in Science and Engineering '13

Nagel, Wolfgang E., Kröner, Dietmar H., Resch, Michael M. (Eds.)
2013, XIII, 697 p. 393 illus., 314 illus. in color.
Transactions of the High Performance Computing Center, Stuttgart (HLRS) 2013.

This book presents the state-of-the-art in simulation on supercomputers. Leading researchers present results achieved on systems of the High Performance Computing Center Stuttgart (HLRS) for the year 2013. The reports cover all fields of computational science and engineering ranging from CFD via computational physics and chemistry to computer science with a special emphasis on industrially relevant applications. Presenting results of one of Europe's leading systems this volume covers a wide variety of applications that deliver a high level of sustained performance. The book covers the main methods in High Performance Computing. Its outstanding results in achieving highest performance for production codes are of particular interest for both the scientist and the engineer. The book comes with a wealth of coloured illustrations and tables of results.

Sustained Simulation Performance 2013



Resch, M.M., Bez, W., Focht, E., Kobayashi, H., Kovalenko, Y. (Eds.)
2013, X, 157 p. 84 illus., 71 illus. in color.

Proceedings of the joint Workshop on Sustained Simulation Performance, University of Stuttgart (HLRS) and Tohoku University, 2013.

This book presents the state-of-the-art in High Performance Computing and simulation on modern supercomputer architectures. It covers trends in hardware and software development in general and specifically the future of high-performance systems and heterogeneous architectures. The application contributions cover Computational Fluid Dynamics, material science, medical applications and climate research. Innovative fields like coupled multi-physics or multi-scale simulations are presented. All papers were chosen from presentations given at the 16th Workshop on Sustained Simulation Performance held in December 2012 at HLRS, University of Stuttgart, Germany and the 17th Workshop on Sustained Simulation Performance at Tohoku University in March 2013.

HLRS Scientific Tutorials and Workshop Report and Outlook

• Rolf Rabenseifner

University of Stuttgart, HLRS

HLRS has installed Hermit, a Cray XE6 system with AMD Interlagos processors and 1 PFlop/s peak performance and will extend with an XC30 system. We strongly encourage you to port your applications to these architectures as early as possible. To support such effort we invite current and future users to participate in the special **Cray XE6/XC30 Optimization Workshops**. With these courses, we will give all necessary information to move applications to this Petaflop system. The Cray XE6 provides our users with a new level of performance. To harvest this potential will require all our efforts. We are looking forward to working with our users on these opportunities. The next four-day course in cooperation with Cray and multi-core optimization specialists is on September 23-26, 2014.

Programming of Cray XK7 clusters with GPUs is taught in **OpenACC Programming for Parallel Accelerated Supercomputers – an alternative to CUDA from Cray perspective** in spring 2015.

These Cray XE6/XC30 and XK7 courses are also presented to the European community in the framework of the **PRACE Advanced Training Centre** (PATC). GCS, i.e., HLRS, LRZ and the Jülich Supercomputer Centre together, serve as one of the first six PATCs in Europe.

One of the flagships of our courses is the week on **Iterative Solvers and Parallelization**. Prof. A. Meister teaches basics and details on Krylov Subspace Methods. Lecturers from HLRS give lessons on distributed memory parallelization with the Message Passing Interface (MPI) and shared memory multi-threading with OpenMP. This course will be presented twice, on September 15-19, 2014 at LRZ and in spring 2015 at HRLS in Stuttgart.

Another highlight is the **Introduction to Computational Fluid Dynamics**. This course was initiated at HLRS by Dr.-Ing. Sabine Roller. She is now a professor at the University of Siegen. It is again scheduled in spring 2014 in Stuttgart and in September/October in Siegen. The emphasis is placed on explicit finite volume methods for the compressible Euler equations. Moreover outlooks on implicit methods, the extension to the Navier-Stokes equations and turbulence modeling are given. Additional topics are classical numerical methods for the solution of the incompressible Navier-Stokes equations, aeroacoustics and high order numerical methods for the solution of systems of partial differential equations.

ISC and SC Tutorials

Georg Hager, Gabriele Jost, Rolf Rabenseifner: **Hybrid Parallel Programming with MPI & OpenMP.** Tutorial 04 at the International Supercomputing Conference, ISC'14, Leipzig, June 22-26, 2014.

Georg Hager, Jan Treibig, Gerhard Wellein: **Node-Level Performance Engineering.** Tutorial 01 at the International Supercomputing Conference, ISC'14, Leipzig, June 22-26, 2014.

Rolf Rabenseifner, Georg Hager, Gabriele Jost: **Hybrid MPI and OpenMP Parallel Programming.** Half-day Tutorial at Super Computing 2013, SC13, Denver, Colorado, USA, November 17-22, 2013.

Our general course on parallelization, the **Parallel Programming Workshop**, October 13-17, 2014 at HLRS, will have three parts: The first two days of this course are dedicated to parallelization with the Message passing interface (MPI). Shared memory multi-threading is taught on the third day, and in the last two days, advanced topics are discussed. This includes MPI-2 functionality, e.g., parallel file I/O and hybrid MPI+OpenMP, as well as the upcoming MPI-3.0. As in all courses, hands-on sessions (in C and Fortran) will allow users to immediately test and understand the parallelization methods. The course language is English.

Several three and four day-courses on **MPI & OpenMP** will be presented at different locations all over the year.

We also continue our series of **Fortran for Scientific Computing** on December 8-12, 2014 and in spring 2015, mainly visited by PhD students from Stuttgart and other universities to learn not only the basics of programming, but also to get an insight on the principles of developing high-performance applications with Fortran.

With **Unified Parallel C (UPC) and Co-Array Fortran (CAF)** in spring 2015, the participants will get an introduction of partitioned global address space (PGAS) languages.

In cooperation with Dr. Georg Hager from the RRZE in Erlangen and Dr. Gabriele Jost from Supersmith, the HLRS also continues with contributions on hybrid MPI & OpenMP programming with tutorials at conferences; see the box on the left page, which includes also a second tutorial with Georg Hager from RRZE.

In the table below, you can find the whole HLRS series of training courses in 2014. They are organized at HLRS and also at several other HPC institutions: LRZ Garching, NIC/ZAM (FZ Jülich), ZIH (TU Dresden), ZDV (Uni Mainz), ZIMT (Uni Siegen) and Uni Oldenburg.

2014 – Workshop Announcements
Scientific Conferences and Workshops at HLRS
13th HLRS/hww Workshop on Scalable Global Parallel File Systems (May 12-14, 2014)
8th ZIH+HLRS Parallel Tools Workshop (date and location not yet fixed)
High Performance Computing in Science and Engineering - The 17th Results and Review Workshop of the HPC Center Stuttgart (September 29-30, 2014)
IDC International HPC User Forum (October 28-29, 2014)
Parallel Programming Workshops: Training in Parallel Programming and CFD
Parallel Programming and Parallel Tools (TU Dresden, ZIH, February 24 - 27)
Cray XE6/XC30 Optimization Workshops (HLRS, March 17-20) (PATC)
Iterative Linear Solvers and Parallelization (HLRS, March 24-28)
Introduction to Computational Fluid Dynamics (HLRS, March 31 - April 4)
GPU Programming using CUDA (HLRS, April 7-9)
Open ACC Programming for Parallel Accelerated Supercomputers (HLRS, April 10-11) (PATC)
Unified Parallel C (UPC) and Co-Array Fortran (CAF) (HLRS, April 14 - 15) (PATC)
Scientific Visualisation (HLRS, April 16-17)
Summer School: Modern Computational Science in Quantum Chemistry (Uni Oldenburg, Aug 25- Sep 05, 2014)
Iterative Linear Solvers and Parallelization (LRZ, Garching, September 15-19)
Introduction to Computational Fluid Dynamics (ZIMT Siegen, September/October)
Message Passing Interface (MPI) for Beginners (HLRS, October 13-14) (PATC)
Shared Memory Parallelization with OpenMP (HLRS, October 15) (PATC)
Advanced Topics in Parallel Programming (HLRS, October 16-17) (PATC)
Parallel Programming with MPI & OpenMP (FZ Jülich, JSC, December 1-3)
Training in Programming Languages at HLRS
Fortran for Scientific Computing (December 8-12) (PATC)
URLs:
http://www.hlrs.de/events/
http://www.hlrs.de/training/course-list/
(PATC): This is a PRACE PATC course

High Performance Computing with Python

Date & Location
June 26-27, 2014
JSC, Forschungszentrum Jülich

Contents
Python is being increasingly used in high-performance computing projects such as GPAW. It can be used either as a high-level interface to existing HPC applications, as embedded interpreter, or directly. This course combines lectures and hands-on session. We will show how Python can be used on parallel architectures and how performance critical parts of the kernel can be optimized using various tools.

Webpage
http://www.fz-juelich.de/ias/jsc/events/hpc-python

Introduction to SuperMUC - the new Petaflop Supercomputer at LRZ

(PATC course)

Date & Location
July 8-11, 2014
LRZ, Garching near Munich

Contents
This four-day workshop gives an introduction to the usage of SuperMUC, the new Petaflop class Supercomputer at LRZ. The first three days are dedicated to presentations by Intel on their software development stack (compilers, tools and libraries); the remaining day will be comprised of talks and exercises delivered by IBM and LRZ on usage of the IBM-specific aspects of the new system (IBM MPI, LoadLeveler, HPC Toolkit) and recommendations on tuning and optimizing for the new system.

Prerequisites
Participants should have good knowledge of HPC-related programming, in particular MPI, OpenMP and at least one of the languages C, C++ or Fortran.

Webpage
http://www.lrz.de/services/compute/courses/

Node-Level Performance Engineering

(PATC course)

Date & Location
July 14-15, 2014
HLRS, Stuttgart

Contents
This course teaches performance engineering approaches on the compute node level. “Performance engineering” as we define it is more than employing tools to identify hotspots and bottlenecks. It is about developing a thorough understanding of the interactions between software and hardware. This process must start at the core, socket, and node level, where the code gets executed that does the actual computational work. Once the architectural requirements of a code are understood and correlated with performance measurements, the potential benefit of optimizations can often be predicted. We introduce a “holistic” node-level performance engineering strategy, apply it to different algorithms from computational science, and also show how an awareness of the performance features of an application may lead to notable reductions in power consumption:

- Introduction
- Practical performance analysis
- Microbenchmarks and the memory hierarchy
- Typical node-level software over heads
- Example problems:
  - The 3D Jacobi solver
  - The Lattice-Boltzmann Method
  - Sparse Matrix-Vector Multiplication
  - Backprojection algorithm for CT reconstruction
- Energy & Parallel Scalability.

Between each module, there is time for Questions and Answers!

Webpage
www.hlrs.de/training/course-list

Industrial Services of the National HPC Centre Stuttgart (HLRS)

Dates & Location
July 16, 2014, and
October 1, 2014
HLRS, Stuttgart

Contents
In order to permanently assure their competitiveness, enterprises and institutions are increasingly forced to deliver highest performance. Powerful computers, among the best in the world, can reliably support them in doing so.
This courses are targeted towards decision makers in companies that would like to learn more about the advantages of using high performance computers in their field of business. They will be given extensive information about the properties and the capabilities of the computers as well as access methods and security aspects. In addition we present our comprehensive service offering - ranging from individual consulting via training courses to visualization. Real world examples will finally allow an interesting insight into our current activities.

Webpage
http://www.hlrs.de/events/

Introduction to Parallel Programming with MPI and OpenMP

Date & Location
August 5-8, 2014
JSC, Forschungszentrum Jülich

Contents
The course provides an introduction to the two most important standards for parallel programming under the distributed and shared memory paradigms: MPI, the Message Passing Interface, and OpenMP. While intended mainly for the JSC guest students, the course is open to other interested persons upon request.

Webpage
http://www.fz-juelich.de/ias/jsc/events/mpi-gsp

GCS – High Performance Computing Courses and Tutorials

Advanced Fortran Topics

Date & Location
September 8-12, 2014
LRZ, Garching near Munich

Contents
This course is targeted at scientists who wish to extend their knowledge of Fortran beyond what is provided in the Fortran 95 standard. Some other tools relevant for software engineering are also discussed. Topics covered include

- object oriented features
- design patterns
- generation and handling of shared libraries
- mixed language programming
- standardized IEEE arithmetic and exceptions
- I/O extensions from Fortran 2003
- parallel programming with coarrays
- source code versioning system (subversion)

To consolidate the lecture material, each day's approximately 4 hours of lectures are complemented by 3 hours of hands-on sessions.

Prerequisites
Course participants should have basic UNIX/Linux knowledge (login with secure shell, shell commands, simple scripts, editor vi or emacs). Good knowledge of the Fortran 95 standard is also necessary, such as covered in the February course at LRZ.

Webpage
http://www.lrz.de/services/compute/courses/

Iterative Linear Solvers and Parallelization

Date & Location
September 15-19, 2014
LRZ, Garching near Munich

Contents
The focus is on iterative and parallel solvers, the parallel programming models MPI and OpenMP, and the parallel middleware PETSc. Different modern Krylov Subspace Methods (CG, GMRES, BiCGSTAB ...) as well as highly efficient preconditioning techniques are presented in the context of real life applications.

Hands-on sessions (in C and Fortran) will allow users to immediately test and understand the basic constructs of iterative solvers, the Message Passing Interface (MPI) and the shared memory directives of OpenMP. This course is organized by LRZ, University of Kassel, HLRS, and IAG.

Webpage
http://www.lrz.de/services/compute/courses/

Cray XE6/XC30 Optimization Workshop

(PATC course)

Date & Location
September 23-26, 2014
HLRS, Stuttgart

Contents
HLRS installed HERMIT, a Cray XE6 system with AMD Interlagos processors and a performance of 1 PFlop/s. Currently, the system is extended by a Cray XC30 system. We invite current and future users to participate in this special course on porting applications to our Cray architectures. The Cray XE6 and Cray XC30 will provide our users with a new level of performance. To harvest this potential will require all our efforts. We are looking forward to working with you on these opportunities. On the first three days, specialists from Cray will support you in your effort porting and optimizing your application on our Cray XE6 / XC30. On the last day, Georg Hager and Jan Treibig from RRZE will present detailed information on optimizing codes on the multicore processors. Course language is English (if required).

Webpage
http://www.hlrs.de/events/

Message Passing Interface (MPI) for Beginners

(PATC course)

Date & Location
October 13-14, 2014
HLRS, Stuttgart

Contents
The course gives a full introduction into MPI-1. Further aspects are domain decomposition, load balancing, and debugging. An MPI-2 overview and the MPI-2 one-sided communication is also taught. Hands-on sessions (in C and Fortran) will allow users to immediately test and understand the basic constructs of the Message Passing Interface (MPI). Course language is english (if required).

Webpage
http://www.hlrs.de/events/

Shared Memory Parallelization with OpenMP

(PATC course)

Date & Location
October 15, 2014
HLRS, Stuttgart

Contents
This course teaches shared memory OpenMP parallelization, the key concept on hyper-threading, dual-core, multi-core, shared memory, and ccNUMA platforms. Hands-on sessions (in C and Fortran) will allow users to immediately test and understand the directives and other interfaces of OpenMP. Tools for performance tuning and debugging are presented. Course language is English (if required).

Webpage
http://www.hlrs.de/events/

Advanced Topics in Parallel Programming

Date & Location
October 16-17, 2014
HLRS, Stuttgart

Contents
Topics are MPI-2 parallel file I/O, hybrid mixed model MPI+OpenMP parallelization, MPI -3.0 parallelization of explicit and implicit solvers and of particle based applications, parallel numerics and libraries, and parallelization with PETSc. Hands-on sessions are included. Course language is English (if required).

Webpage
http://www.hlrs.de/events/

GCS – High Performance Computing Courses and Tutorials

Scientific Visualization

Date & Location
October 20-21, 2014
HLRS, Stuttgart

Contents
This two day course is targeted at researchers with basic knowledge in numerical simulation, who would like to learn how to visualize their simulation results on the desktop but also in Augmented Reality and Virtual Environments. It will start with a short overview of scientific visualization, following a hands-on introduction to 3D desktop visualization with COVISE. On the second day, we will discuss how to build interactive 3D Models for Virtual Environments and how to set up an Augmented Reality visualization.

Webpage
http://www.hlrs.de/events/

GPU Programming using CUDA

Date & Location
October 22-24, 2014
HLRS, Stuttgart

Contents
The course provides an introduction to the programming language CUDA, which is used to write fast numeric algorithms for NVIDIA graphics processors (GPUs). Focus is on the basic usage of the language, the exploitation of the most important features of the device (massive parallel computation, shared memory, texture memory) and efficient usage of the hardware to maximize performance. An overview of the available development tools and the advanced features of the language is given.

Webpage
http://www.hlrs.de/events/

Data Analysis and Data Mining with Python

Date & Location
November 10-12, 2014
JSC, Forschungszentrum Jülich

Contents
Pandas, matplotlib, and scikit-learn make Python a powerful tool for data analysis, data mining, and visualization. All of these packages and many more can be combined with IPython to provide an interactive extensible environment. In this course, we will explore matplotlib for visualization, pandas for time series analysis, and scikit-learn for data mining. We will use IPython to show how these and other tools can be used to facilitate interactive data analysis and exploration.

Webpage
http://www.fz-juelich.de/ias/jsc/events/python-data-mining

Introduction to the Programming and Usage of the Supercomputer Resources at Jülich

Date & Location
November 27-28, 2014
JSC, Forschungszentrum Jülich

Contents
This course gives an overview of the supercomputers at Jülich. Especially new users will learn how to program and use these systems efficiently. Topics discussed are: system architecture, usage model, compilers, tools, monitoring, MPI, OpenMP, performance optimization, mathematical software, and application software.

Webpage
http://www.fz-juelich.de/ias/jsc/events/sc-nov

Parallel Programming with MPI and OpenMP

Date & Location
December 1-3, 2014
JSC, Forschungszentrum Jülich

Contents
The focus is on programming models MPI and OpenMP. Hands-on sessions (in C and Fortran) will allow users to immediately test and understand the basic constructs of the Message Passing Interface (MPI) and the shared memory directives of OpenMP. Course language is German. This course is organized by JSC in collaboration with HLRS. Presented by Dr. Rolf Rabenseifner, HLRS.

Webpage
http://www.fz-juelich.de/ias/jsc/events/mpi

Fortran for Scientific Computing

Date & Location
December 8-12, 2014
HLRS, Stuttgart

Contents
This course is dedicated for scientists and students to learn (sequential) programming scientific applications with Fortran. The course teaches newest Fortran standards. Hands-on sessions will allow users to immediately test and understand the language constructs.

Webpage
http://www.hlrs.de/events/

Contents

If you would like to receive inside regularly, send an email with your postal address to klank@hlrs.de